Rita De Siano
Luca Pennacchio
(Eds)

Annals of CRISEI 2024

Sources of Inequality and Public Policies

Social Conflicts, Insurance and Labour Markets

CRISEI

ISBN: 978-88-98279-15-9

Contents

Contributors

Giulia Bettin, Università Politecnica of Marche

Mariangela Bonasia, University of Naples Parthenope

Maria Rosaria Carillo, University of Naples Parthenope

Salvatore Carrozzo, University of Naples Parthenope

Valentina Chiariello, University of Naples Parthenope

Alessandro De Iudicibus, University of Campania Luigi Vanvitelli

Giuseppe De Marco, University of Naples Parthenope

Rita De Siano, University of Naples Parthenope

Alfredo Del Monte, University of Naples Federico II

Alessandra Di Pietro, University of Turin

Christian Di Pietro, University of Naples Parthenope

Vincenzo Lombardo, University of Naples Parthenope

Sara Moccia, University of Naples Federico II

Luca Pennacchio, University of Naples Parthenope

Mariafortuna Pietroluongo, University of Naples Parthenope

Emma Sarno, University of Naples L'Orientale

Taslim Uddin, Asian University for Woman

Alberto Zazzaro, University of Naples Federico II

Introduction

Rita De Siano and Luca Pennacchio

This volume collects a set of recent contributions developed within the research activities of CRISEI, addressing a wide range of topics related to economic development, institutional dynamics, and market functioning. Despite the diversity of approaches and fields of application, the papers share a common focus on the role of public policies and institutional and behavioural factors in shaping economic outcomes. In particular, the contributions explore how constraints—whether financial, institutional, or informational—interact with individual and collective decisions, ultimately influencing efficiency, inequality, and growth.

A first group of contributions focuses on the determinants of economic development and social mobility. The paper by Carillo, Lombardo, and Zazzaro investigates the interplay between entrepreneurial human capital, family connections, and intergenerational mobility within a dynastic framework. By introducing financial constraints on access to entrepreneurial education, the authors show that talent allocation is shaped not only by ability but also by inherited resources. Their model highlights the emergence of different long-run equilibria: one characterized by partial mobility and efficient selection, and another marked by persistent inequality and low growth. The analysis provides important insights into how credit constraints and family structures may generate inefficient persistence in economic positions, limiting both social mobility and aggregate performance.

A related perspective on development is offered by Carillo, Chiariello, and De Siano, who examine whether ethnic fractionalization plays a role in the controversial aid-corruption relationship that characterizes most recipient countries worldwide. The hypothesis is that ethnic fractionalization may exert pressure on the redistributive process within a given society, thereby clarifying why aid can sometimes be diverted from more productive uses, causing it to be ineffective or even harmful. By applying an appropriate quantile regression approach to a panel dataset including 38 African countries, the authors show that foreign aid may exert a direct beneficial effect on corruption only in the least corrupt countries and the reverse in more corrupt countries. Ethnic fractionalization, instead, exerts a direct, strong, detrimental effect on corruption but also modifies the effectiveness of aid campaigns depending on the corruption level, raising doubts about the adequacy of equal management of international aid worldwide.

A second group of papers investigates the functioning of markets and public policies, with particular attention to efficiency and institutional quality. The contribution by Del Monte, De

Iudicibus, Moccia, and Pennacchio analyzes the determinants of the speed of public spending, focusing on the role of institutional quality, administrative capacity, and government decentralization. Using detailed data on public projects implemented in Italy, the authors show that both institutional quality and administrative capacity significantly affect the efficiency of resource allocation. Moreover, their findings suggest that decentralization may have ambiguous effects, depending on the quality of local institutions. This highlights the importance of governance structures in determining the effectiveness of public investment and the broader impact of cohesion policies.

The theme of the effectiveness of public resources is also detected in the paper, "Financial crisis and the convergence of European welfare provision" by Mariangela Bonasia and Rita De Siano. To this aim, the paper seeks to evaluate whether and to what extent the recent global financial crisis and its economic aftermaths affected the dynamics of national welfare provisions in European countries. Based on data from 16 Western countries covering the period 1990-2013, the analysis examines the behaviour of per capita levels of total social expenditure and its main functions: old-age, survivors, incapacity-related, labour, and healthcare. The empirical analysis reveals the presence of a strong conditional convergence process of per capita total public social expenditure and that devolved to functions attracting most of the social resources. The findings reveal that the 2007 financial crisis did not change the old age-survivors-incapacity and labour policies' convergence trend, while it contributed to increasing differences among national indicators of total social provision and the health sector, with the latter showing a turnaround relative to the period before the crisis. Healthcare is confirmed to be a less productive sector and therefore the one that may be cut in the presence of resource availability constraints. This proved to be a critical point during the COVID-19 in 2020, as reduced capacity limited governments' ability to respond effectively

Market dynamics and strategic interactions are at the core of the contribution by De Marco, Di Pietro, and Pietroluongo, which studies equilibrium outcomes in the non-life insurance market. The authors develop a theoretical framework in which insurers set premiums under uncertainty, accounting for both competitive and collusive behaviour. By introducing the concept of collusive and Kantian equilibria, the paper provides a novel perspective on how cooperation among firms may emerge and how it affects market efficiency. The analysis also highlights the role of regulation in shaping market outcomes and in detecting potential collusive behaviour.

Carrozzo and Di Pietro address labour market dynamics by examining the degree of substitutability between older and younger workers. Using Italian administrative data and a novel identification strategy, the authors find that the two groups are imperfect substitutes in production. This result has important policy implications, suggesting that increases in older workers' employment do not necessarily crowd out younger workers. Instead, labor market outcomes depend on more complex interactions between demographic changes, institutional reforms, and production structures.

Finally, Bettin, Sarno, Uddin and Zazzaro explore how women's economic empowerment within households affects financial portfolio choices. Considering Italy between 2004 and 2020, the authors find that a higher female income share is associated with lower exposure to risky assets in male-headed households and with increased risky investments in female-headed households. This evidence supports a unitary decision-making framework à la Becker, in which the identity and confidence of the household financial decision-maker are central. The main mechanism underlying household portfolio choices is self-confidence, because the household head's financial-sector employment mitigates these effects, whereas similarity in partners' risk preferences is not important.

Taken together, the contributions in this volume provide a multifaceted view of the mechanisms underlying economic performance and social outcomes. They highlight the importance of institutions, public policies, and strategic interactions in shaping both micro-level decisions and macro-level dynamics. By combining theoretical models with empirical analyses, the papers offer valuable insights for both academic research and policy design, contributing to a deeper understanding of the challenges facing contemporary economies.

Entrepreneurial human capital, family connections and social mobility*

Maria Rosaria Carillo
University of Naples Parthenope
mr.carillo@uniparthenope.it

Vincenzo Lombardo
University of Naples Parthenope
vincenzo.lombardo@uniparthenope.it

Alberto Zazzaro
University of Naples Federico II
alberto.zazzaro@unina.it

June 15, 2024

Abstract

This paper extends Carillo et al. [2019] by introducing bequest-financed monetary costs of entrepreneurial education into a dynastic family-firm model with heterogeneous talent. Family firms can be managed through two technologies: entrepreneurial human capital or inherited family-specific managerial assets. Under credit frictions, educational access depends on parental resources, so occupational choice is jointly determined by descendants' ability and dynastic wealth. The model delivers two long-run regimes. First, when education costs and inherited-asset productivity are moderate, the economy converges to a polarization equilibrium with coexistence of well-managed and badly managed family firms, positive social mobility, and positive growth. Second, when worker dynasties are systematically constrained, the economy falls into a low-mobility equilibrium with dynastic entrepreneurial persistence and lower growth. The framework identifies how education finance, family connections, and managerial selection interact to shape social mobility, talent allocation, and development.

Keywords: family firms, entrepreneurial human capital, social mobility, wealth constraints, growth

JEL codes: D31, J62, L26, O33, O40

*Maria Rosaria Carillo and Vincenzo Lombardo greatly acknowldge research funds received by the Ministry of University and Research of Italian Government under the PRIN project "The Role of Skill-Biased Technological Change in Immiserizing Growth, Spatial Polarization, and Product Quality Upgrading" (Prin-2022-Carillo, ID project 2022ZKA4PZ, CUP I53D23002850006). The views and opinions expressed are solely those of the author and do not necessarily reflect those of the Ministry of University and Research, and the Italian Government cannot be held responsible for them.

1 Introduction

Family firms are one of the most persistent organizational forms across countries and development stages. Their persistence is often associated with governance advantages inside kinship networks, stronger internal monitoring, and longer planning horizons [Burkart et al., 2003, Bertrand and Schoar, 2006]. At the same time, dynastic control can reduce the quality of managerial selection when succession is primarily driven by lineage rather than competence [Bennedsen et al., 2007]. This fundamental tension between continuity and meritocratic allocation is central to understanding why family capitalism is growth-enhancing in some contexts and growth-reducing in others.

The empirical evidence is now clear on one point: family ownership by itself is not the key object. What matters is the quality of management and the institutional environment in which succession decisions are taken [Bloom and Van Reenen, 2007, 2010, Alesina and Giuliano, 2015]. In some economies, family firms led by highly educated and professionally selected heirs perform comparably to non-family firms. In others, firms remain under dynastic control even when managerial quality is low, and this correlates with weaker innovation and lower productivity growth [Bennedsen and Foss, 2015]. The relevant margin is therefore not "family versus non-family", but "ability-based versus inheritance-based managerial assignment".

This paper builds a formal theory of that margin. We study how entrepreneurial human capital, inherited family connections, and education finance interact to determine occupational choice, social mobility, and growth. The paper is an extension of Carillo et al. [2019], published in the *Journal of Economic Growth*. Relative to that framework, we introduce a monetary cost of entrepreneurial education financed through intergenerational bequests under credit frictions. This addition is small in appearance, but it changes the selection mechanism substantially: the ability to become an entrepreneur now depends jointly on descendants' talent and parental resources.

The extension is motivated by a broad set of fixed costs faced by potential entrepreneurs: formal education, technical training, certification, managerial mentoring, and setup costs that are hard to collateralize. Following the logic in Galor and Zeira [1993], we assume that these costs cannot be debt-financed. Descendants can invest in entrepreneurial human capital only if parental bequests are high enough. As a result, educational access becomes dynastic and endogenous to past occupational outcomes.

Relative to the baseline structure in Carillo et al. [2019], this modification changes both the individual and aggregate content of the model. At the individual level, occupational choice now includes a feasibility layer in addition to a profitability layer: descendants may prefer entrepreneurial education but be unable to finance it. At the aggregate level, growth depends on how feasibility and profitability interact through wages and bequests, so the composition of entrepreneurs becomes jointly determined by talent, inherited ties, and intergenerational liquidity.

This distinction is central for interpretation. In the original framework, talented descendants can upgrade management whenever returns justify it. In the present extension, the same descendants can be blocked by financing constraints inherited from parental occupational history. The resulting inefficiency is dynamic: talent is observable and potentially productive, but it is not implemented because dynastic resources and educational access are jointly determined by past allocation outcomes.

This channel creates a dynamic wedge between private and social efficiency. From a private perspective, low-ability heirs may rationally keep family control by relying on inherited relational assets. From a social perspective, this may block high-ability entrants from worker dynasties when educational finance is tight. The model captures both effects in a unified environment and quantifies the equilibrium consequences for firm composition and long-run growth.

Three economic ingredients are central. The first is a dual managerial technology: entrepreneurial dynasties can run firms either through acquired entrepreneurial human capital, which is ability-sensitive and costly, or through family-specific assets, which are lineage-sensitive and do not require direct monetary educational expenditure. The second is intergenerational financing: entrepreneurial education requires a fixed monetary payment proportional to the technology frontier and must be financed through parental bequests. The third is a general-equilibrium feedback: wages affect current occupational choices and future educational feasibility by shaping the bequest process.

These ingredients jointly generate nonlinear dynastic sorting. Ability alone does not pin down occupation. The same talented descendant can become an entrepreneur in one dynasty and a worker in another because the financing constraint depends on parental income and parental management mode. Conversely, low-ability heirs may remain entrepreneurs when family assets are productive enough to insure dynastic continuity.

The framework allows us to answer two broad questions. First, when does the economy sustain a *polarization equilibrium* in which well-managed and badly managed family firms coexist, with positive but incomplete social mobility? Second, under what conditions does the economy converge to a *low-mobility trap* where worker dynasties cannot finance entrepreneurial entry and entrepreneurial status becomes largely inherited?

The model produces tractable threshold rules. For descendants of entrepreneurs, we derive a talent threshold above which educated management dominates crony management, wage thresholds that govern switching between entrepreneurship and paid work, and a parental-ability threshold that determines education feasibility for descendants. For worker dynasties, entrepreneurial entry is feasible only above a bequest-implied parental-ability cutoff. These thresholds fully characterize occupational regions and permit clear comparative statics.

The first main result is the existence of a stationary equilibrium with positive entry and exit, positive growth, and endogenous firm-type polarization, provided that education costs and inherited-asset productivity are not too high. In this equilibrium, descendants of crony entrepreneurs are always constrained; by contrast, a positive mass of workers' descendants and descendants of entrepreneurially educated entrepreneurs can finance education. Social mobility is therefore strictly positive but limited.

The second main result is the existence of a distinct stationary equilibrium with no worker-to-entrepreneur mobility and lower growth. In that regime, workers' dynasties systematically fail to finance entrepreneurial education. Entrepreneurship survives only through incumbent dynasties, where descendants choose between crony and educated management inside existing family firms. Growth remains positive but falls relative to the polarization regime because selection from outside dynasties is shut down.

Comparative statics are sharp and economically intuitive. Higher education costs increase the constrained mass, reduce the share of firms managed through entrepreneurial human capital, and lower growth. A higher productivity of inherited family assets increases the private value of dynastic continuity and weakens meritocratic reallocation, reducing both social mobility and long-run efficiency. In short, family connections can be productive in weak institutional environments, but they can become a source of dynamic misallocation when they substitute for entrepreneurial investment.

The paper contributes to three literatures. First, it complements models of talent allocation and dynastic management [Caselli and Gennaioli, 2013, Baumol, 1990] by introducing an explicit interaction between managerial technology and educational finance. Second, it contributes to the macro-development literature on inequality and mobility [Galor and Zeira, 1993, Galor and Tsiddon, 1997, Galor and Moav, 2000] by embedding credit constraints inside a dynastic family-firm environment. Third, it adds to the economics of family business [Burkart et al., 2003, Bertrand and Schoar, 2006, Bennedsen et al., 2007, Bennedsen and Foss, 2015] by formalizing when family continuity supports growth and when it locks economies into low-mobility equilibria.

The model also provides a clear policy map. Interventions that reduce private education costs, broaden household liquidity, or increase the return to merit-based succession can jointly improve talent allocation and growth. The key point is complementarity: education policy without governance reform may expand supply of skilled potential entrants but preserve low-merit incumbency rents; governance reform without educational access may improve selection among incumbents but leave talented outsiders constrained. The strongest dynamic effects arise when both margins are addressed.

Although the analysis is positive, the welfare interpretation is immediate. Whenever high-ability constrained descendants are pushed into wage work while low-ability heirs retain entrepreneurial control through family assets, equilibrium departs from a productivity-based assignment benchmark. The loss has a static component (lower current output) and a dynamic component (slower accumulation of entrepreneurial human capital and frontier growth).

The framework also helps organize several empirical regularities that are often studied separately. Many countries show persistent family ownership together with wide managerial-quality dispersion inside family firms. Upward mobility into entrepreneurship is often concentrated among descendants from better-off worker households rather than evenly distributed across income groups. Policy episodes that expand educational access tend to be more growth-enhancing when accompanied by governance reforms affecting succession and control rights. The model provides a unified interpretation of these facts through interacting thresholds.

A second contribution is measurement discipline. The theory implies that the aggregate family-firm share is not a sufficient statistic for development performance. Similar ownership shares can arise in a high-mobility equilibrium with active merit-based upgrading or in a low-mobility equilibrium sustained by dynastic exclusion. Distinguishing between these cases requires joint evidence on intergenerational transitions, within-family management dispersion, and the share of educated managers among entrepreneurs.

The extension also clarifies why reform effects can appear delayed. Because education is financed through bequests, policy changes alter future feasible sets through an intergenerational state variable, not only through current profits. Immediate effects on observed mobility can therefore be modest even when medium-run effects are large once the bequest distribution adjusts. This prediction is useful for empirical design because it cautions against evaluation windows that are too short to capture dynastic propagation.

The rest of the paper is organized as follows. Section 2 presents the environment and the bequest-financed education mechanism. Section 3 derives occupational choices and threshold conditions. Section 4 characterizes steady-state equilibria, transition mechanisms, analytical extensions, and quantitative implications. Section 7 concludes.

2 Economic Environment

2.1 Demography, abilities, and dynasties

Time is discrete, $t = 0, 1, 2, \ldots$. In each period a unit mass of descendants is born. Individuals live for two periods. In childhood they choose educational track; in adulthood they choose occupation, produce, consume, and leave bequests.

Each descendant i draws innate ability $a_t^i \sim U[0, 1]$. Ability is iid across generations. Persistence in occupations therefore comes from dynasty-specific resources and occupational inheritance, not from inherited talent in the model.

There are two adult occupations: worker (ω) and entrepreneur (e). Workers supply efficiency labor and receive

$$I_{w,t+1}^{i,\omega} = w_{t+1} a_{t+1}^i, \tag{1}$$

where w_{t+1} is the competitive wage. Hence, in the baseline normalization, general human capital in wage employment is

$$h_{t+1}^{i,\omega} = a_{t+1}^i. \tag{2}$$

2.2 Managerial capital

Following Carillo et al. [2019], managerial capital is a composite of acquired entrepreneurial skills and inherited family-specific assets. For an entrepreneur i at $t + 1$:

$$m_{t+1}^i = \tau_{t+1}^i a_{t+1}^i + \iota_{t+1}^i (1 - \tau_{t+1}^i) \Phi_{t+1}, \tag{3}$$

where $\tau_{t+1}^i \in \{0, 1\}$ indicates management mode (entrepreneurial human-capital management if $\tau = 1$, family-asset management if $\tau = 0$), $\iota_{t+1}^i \in \{0, 1\}$ indicates whether the descendant belongs to an entrepreneurial dynasty (only in that case family assets are available), $\Phi_{t+1} \equiv \phi(1 - g_{t+1})$ is the effective productivity of inherited family assets, and $g_{t+1} \equiv (A_{t+1} - A_t)/A_t$ is frontier growth.

Equation (3) makes the model transparent. Entrepreneurial human capital is ability-based and requires educational investment; family-asset management is lineage-based and does not require monetary educational expenditure. The parameter ϕ captures inherited reputation, relational contracts, local trust, and dynastic business ties.

2.3 Production structure and reduced-form profits

Each entrepreneur runs one firm with technology

$$y_{t+1}^{i} = A_{t+1} m_{t+1}^{i} \left(h_{t+1}^{i}\right)^{1-\alpha}, \qquad \alpha \in (0,1), \tag{4}$$

where h_{t+1}^{i} is labor hired competitively at wage w_{t+1}. Static profit maximization yields labor demand:

$$h_{t+1}^{i} = \left(\frac{(1-\alpha) A_{t+1} m_{t+1}^{i}}{w_{t+1}}\right)^{1/\alpha}. \tag{5}$$

Substituting (5) into profits gives a reduced form:

$$\pi_{t+1}^{i} = \Psi_{t+1} \left(m_{t+1}^{i}\right)^{1/\alpha}, \qquad \Psi_{t+1} \equiv \left(\frac{\theta A_{t+1}}{w_{t+1}^{1-\alpha}}\right)^{1/\alpha}, \tag{6}$$

where $\theta > 0$ absorbs constants from the static optimization problem.

Using (3) inside (6), entrepreneurial income is

$$I_{e,t+1}^{i,e} = \begin{cases} \Psi_{t+1} \left[\phi(1-g_{t+1})\right]^{1/\alpha} & \text{if } \tau_{t+1}^{i} = 0 \text{ (crony management)}, \\ \Psi_{t+1} \left(a_{t+1}^{i}\right)^{1/\alpha} - s_t & \text{if } \tau_{t+1}^{i} = 1 \text{ (educated management)}, \end{cases} \tag{7}$$

where $s_t = s_0 A_t$ is the private monetary cost of entrepreneurial education.

The structure in (3)–(7) clarifies the economic logic behind occupational sorting. Descendants of entrepreneurs can always choose family-asset management if they stay in business; all descendants can choose educated management only if they can finance education.

2.4 Wage determination and equilibrium feedback

Let N_t be the mass of entrepreneurs and μ_t the distribution of managerial quality. Labor-market clearing can be represented by

$$w_t = \mathcal{W}(A_t, N_t, \mu_t), \tag{8}$$

with $\partial \mathcal{W}/\partial A_t > 0$ and $\partial \mathcal{W}/\partial \mu_t > 0$. Better managerial selection raises labor demand and wages; a compositional shift toward low-quality crony management weakens wage growth.

This feedback is crucial: wages influence current occupational payoffs and, through bequests, next-generation education feasibility. Hence policies or shocks that initially affect managerial composition can later affect social mobility through general-equilibrium price adjustments.

2.5 Bequests and education financing

As in Galor and Zeira [1993], borrowing for education is unavailable. Education costs can only be financed through parental bequests. Parents choose consumption and bequests during adulthood.

An adult with wealth

$$\Omega_t^{i} = I_t^{i} - \iota s_{t-1} + b_{t-1}^{i}, \tag{9}$$

solves

$$\max_{c_t^i, b_t^i} \; u_t^i = \gamma \ln c_t^i + (1-\gamma) \ln b_t^i \quad \text{s.t.} \quad c_t^i + b_t^i = \Omega_t^i, \tag{10}$$

where $\iota = 1$ if the parent invested in entrepreneurial education and $\iota = 0$ otherwise. The optimal bequest is

$$b_t^{i*} = (1-\gamma)\left(I_t^i - \iota s_{t-1} + b_{t-1}^i\right). \tag{11}$$

Along the invariant limiting dynasty distribution, bequests are proportional to net income:

$$\hat{b}_t^i = \lambda\left(I_t^i - \iota s_{t-1}\right), \qquad \lambda \equiv \frac{1-\gamma}{\gamma}. \tag{12}$$

Descendants can choose entrepreneurial education only if $\hat{b}_t^i \geq s_t$.

2.6 Timing and information structure

Each period unfolds in four stages. In stage 1, technology A_t and inherited bequests are observed. In stage 2, descendants choose educational track subject to the financing constraint. In stage 3, adult occupational choices are made after wage expectations are formed. In stage 4, incomes are realized, consumption occurs, and bequests for the next generation are set.

This timing implies that wealth constraints operate ex ante, before descendants know the full realization of equilibrium wages in adulthood. The model therefore captures a realistic source of risk: constrained descendants may be prevented from taking educational investments that would have had high expected returns ex post. Dynastic inequality affects not only deterministic opportunities but also exposure to upside risk.

Information is complete at the aggregate level but family-specific at the dynastic level. Each dynasty knows parental ability, parental occupation, and inherited wealth, and takes wage schedules as given. Because ability is iid across generations, persistence in occupational outcomes does not come from genetic transmission of talent in the model; it comes from financial transmission and occupational inheritance. This feature makes the social-mobility mechanism transparent and separates it from biological channels.

The bequest recursion in (11) creates a forward-looking channel even under log utility. A dynasty that chooses low-return management today can reduce descendants' feasible opportunity set tomorrow, and vice versa. Thus, managerial choices have a dynastic shadow price through their effect on next-generation education access.

From a macro perspective, the mapping from period-t occupational distribution to period-$(t+1)$ feasible sets is the key state transition of the economy. Policy reforms that change education costs or governance incentives alter this mapping directly; reforms that only affect static profitability alter it indirectly through wages and bequests.

3 Occupational Choice with Wealth Constraints

3.1 Income thresholds

For descendants of entrepreneurs, compare the two entrepreneurial technologies in (7). The ability threshold above which educated management dominates crony management is

$$\bar{a}^s_{t+1} = \left\{ [\phi(1-g_{t+1})]^{1/\alpha} + \frac{s_0}{1+g_{t+1}} \left(\frac{w_{t+1}}{\theta A_{t+1}} \right)^{\frac{1-\alpha}{\alpha}} \right\}^{\alpha}. \tag{13}$$

Comparing entrepreneurial income and wage income yields two additional thresholds:

$$a^{\phi,s}_{t+1} = \left(\frac{\theta A_{t+1}\phi(1-g_{t+1})}{w_{t+1}} \right)^{1/\alpha}, \tag{14}$$

$$\left(\frac{\theta A_{t+1}}{w^{1-\alpha}_{t+1}} a^{a,s}_{t+1} \right)^{1/\alpha} = w_{t+1}a^{a,s}_{t+1} + s_t. \tag{15}$$

For workers' descendants, the entry threshold into entrepreneurship is also $a^{a,s}_{t+1}$ because they can only use educated management.

Define wage thresholds $\hat{w}^s_{t+1}$ and $\tilde{w}^s_{t+1}$ such that

$$a^{\phi,s}_{t+1} = a^{a,s}_{t+1} = \bar{a}^s_{t+1} \quad \text{at } w_{t+1} = \hat{w}^s_{t+1}, \tag{16}$$

and

$$a^{a,s}_{t+1} = 1 \quad \text{at } w_{t+1} = \tilde{w}^s_{t+1}. \tag{17}$$

Lemma 1. *$\hat{w}^s_{t+1} \le \tilde{w}^s_{t+1}$ if $\phi \le \bar{\phi}^s$, where $\bar{\phi}^s$ is implicitly defined by*

$$s_0(\bar{\phi}^s)^{(1-\alpha)/\alpha} = \theta \left[1 - (\bar{\phi}^s)^{1/\alpha} \right]. \tag{18}$$

Lemma 1 has a direct economic meaning. The ordering $\hat{w}^s_{t+1} \le \tilde{w}^s_{t+1}$ implies that, over a non-empty wage interval, high-ability entrepreneurial descendants still find it optimal to remain entrepreneurs *and* adopt educated management rather than either switching to wage work or relying on inherited ties. The condition $\phi \le \bar{\phi}^s$ is therefore a tractability and economics condition: family assets can be valuable, but not so valuable that they dominate talent-based management for almost every relevant wage.

In plain language, the lemma identifies when family control remains "open" to internal upgrading. If inherited rents are too strong, the dynasty has weak incentives to invest in managerial capability; low-merit continuity becomes privately attractive, and talented heirs may even prefer wage careers. If inherited rents are moderate, dynastic firms still have an internal meritocratic margin: strong heirs upgrade management from crony to entrepreneurial human capital.

3.2 Regime partition in ability space

Occupational outcomes can be interpreted as a sequence of filters in ability space. First, feasibility: descendants must satisfy the bequest condition to access entrepreneurial education. Second, technological choice: conditional on entrepreneurship, descendants compare educated and crony management at threshold $\bar{a}^s_{t+1}$. Third, occupation choice: descendants compare entrepreneurial and wage payoffs through $a^{\phi,s}_{t+1}$ and $a^{a,s}_{t+1}$.

This decomposition clarifies why family firms affect both extensive and intensive margins. On the extensive margin, family background changes who can become an entrepreneur. On the intensive margin, inherited assets change how incumbent firms are run. When financing constraints are mild, the model behaves mainly as a talent-allocation model with family insurance in low-ability states. When financing constraints are strong, parental occupation and parental ability dominate own talent in determining access to entrepreneurial careers.

The intermediate case is empirically the most relevant: coexistence of constrained and unconstrained dynasties generates a dispersed quality distribution among family firms, with high-ability educated managers and low-ability crony managers operating side by side. This coexistence is a central source of the polarization equilibrium characterized below.

3.3 Bequest-based wealth constraints

Using (12), bequests of entrepreneurs' dynasties are

$$\hat{b}^i_t = \begin{cases} \lambda \left[\dfrac{\phi(1-g_t)\theta A_t}{w_t^{1-\alpha}}\right]^{1/\alpha}, & a^i_t < \bar{a}^s_t, \\ \lambda \left[\left(\dfrac{a^i_t \theta A_t}{w_t^{1-\alpha}}\right)^{1/\alpha} - s_{t-1}\right], & a^i_t \geq \bar{a}^s_t \text{ and } \hat{b}^i_{t-1} \geq s_{t-1}. \end{cases} \tag{19}$$

Define $\hat{a}^s_t$ as the parental-ability threshold above which descendants can finance entrepreneurial education:

$$\hat{a}^s_t = \left\{ s_0 \left[\frac{1+\gamma g_t}{(1-\gamma)(1+g_t)}\right] \left(\frac{w_t}{\theta A_t}\right)^{\frac{1-\alpha}{\alpha}} \right\}^{\alpha}. \tag{20}$$

Proposition 1 (Wealth constraints in entrepreneurial dynasties)**.** *The descendants of entrepreneurs are wealth-constrained according to the position of $\hat{a}^s_t$ relative to $\bar{a}^s_t$ and $a^{a,s}_t$:*

(a) if $\hat{a}^s_t \leq \bar{a}^s_t$, no descendants are constrained;

(b) if $\bar{a}^s_t < \hat{a}^s_t \leq a^{a,s}_t$, only descendants of crony parents are constrained;

(c) if $a^{a,s}_t < \hat{a}^s_t \leq 1$, all descendants of crony parents and a fraction of descendants of educated parents are constrained;

(d) if $\hat{a}^s_t > 1$, all descendants are constrained.

Proposition 1 clarifies that "being an entrepreneurial dynasty" is not enough to guarantee intergenerational access to entrepreneurial human capital. What matters is *which type* of entrepreneur the parent is. Crony management produces flatter bequests and therefore weaker educational access for descendants. Educated management produces bequests increasing in ability and can sustain a self-reinforcing high-skill branch of dynasties.

This classification is central for interpretation. Case (b) is the empirically relevant "middle" regime in which both mobility and persistence coexist: descendants of educated entrepreneurs can keep investing, while descendants of crony entrepreneurs are locked into low-investment trajectories. Case (c) is even more restrictive because constraints spill over to part of educated dynasties. Case (d) is a full dynastic compression where education finance collapses as a selection device.

An additional implication is that dynastic persistence is internally heterogeneous. Two entrepreneurial lineages can display similar continuity in ownership but diverge sharply in managerial quality and long-run contribution to growth. Lineages repeatedly financing entrepreneurial human capital move toward a high-productivity branch with stronger bequests. Lineages repeatedly relying on crony management remain fragile, more exposed to wage competition, and less able to transmit feasible opportunity sets to descendants.

From (13) and (20), define two wage thresholds:

$$\hat{w}_t^s = \theta A_t \left\{ \phi(1 - g_{t+1}) \left[\frac{1-\gamma}{s_0 \gamma} \right]^{\alpha} \right\}^{1/(1-\alpha)}, \tag{21}$$

$$\tilde{w}_t^s = \theta A_t \left[\frac{(1-\gamma)(1+g_t)}{s_0(1+\gamma g_t)} \right]^{\alpha/(1-\alpha)}. \tag{22}$$

Lemma 2. *If $\phi \leq \gamma^{\alpha}$, then $\hat{w}_t^s \leq \tilde{w}_t^s$.*

Lemma 2 links the productivity of inherited ties (ϕ) to the distribution of financing constraints across entrepreneurial lineages. When ϕ is sufficiently low relative to saving behavior, descendants of crony parents are more likely to be constrained than descendants of educated parents. In other words, dynastic persistence is not uniform: some family lines remain financially fragile even inside the entrepreneurial class.

Economically, this result prevents an extreme and implausible inversion in which low-quality dynastic management systematically generates larger educational opportunities than high-quality entrepreneurial management. The lemma therefore supports a coherent ranking of dynastic trajectories and is a key building block for Propositions 2 and 4.

3.4 Occupational choice rules

Proposition 2 (Descendants of entrepreneurs)**.** *If $\phi \leq \min\{\bar{\phi}^s, \gamma^{\alpha}\}$, occupational choices satisfy:*

(a) if $w_t \leq \hat{w}_t^s$, no wealth constraints bind; choices are governed by talent thresholds in (13)–(15);

(b) if $\hat{w}_t^s < w_t \leq \tilde{w}_t^s$, constrained descendants cannot choose educated management and select between crony management and wage work according to $a_{t+1}^{\phi,s}$;

(c) if $w_t > \tilde{w}_t^s$, all descendants are constrained; only those with $a_{t+1}^i < a_{t+1}^{\phi,s}$ keep the family firm.

Proposition 2 provides a clean regime map for entrepreneurial dynasties. Panel (a) is the "full-choice" region: descendants compare all occupations and technologies, so sorting is mostly ability-driven. Panel (b) is the "partial-exclusion" region: constrained descendants lose access to educated management and choose between crony continuity and wage work. Panel (c) is the "full-exclusion" region for entrepreneurial education, where dynastic continuity survives only for descendants below the crony-versus-wage cutoff.

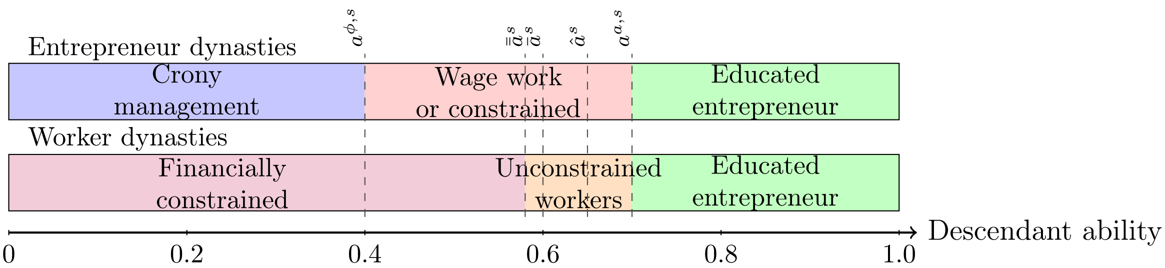


Figure 1: Threshold map and occupational sorting under a stylized numerical ordering of cutoffs. The diagram shows financing thresholds ($\hat{a}^s$, $\bar{\bar{a}}^s$), the internal management threshold ($\bar{a}^s$), and occupational thresholds ($a^{\phi,s}$, $a^{a,s}$).

The practical implication is that higher wages need not increase entrepreneurship among entrepreneurial dynasties. When wages rise beyond financing thresholds, high-ability constrained descendants can optimally leave family firms and become workers. This is why the composition of entrepreneurs may worsen even when aggregate labor income rises.

The regime map also yields testable implications for succession data. In partial-exclusion regimes, one should observe high-ability descendants from constrained entrepreneurial dynasties disproportionately selecting wage careers, while lower-ability descendants preserve control through inherited ties. In full-choice regimes, managerial upgrading should be more tightly aligned with descendant ability. Observing how these patterns change after education-finance reforms provides a direct way to test whether constraints or preferences drive dynastic continuity.

3.5 Threshold comparative statics in the core model

The central thresholds admit transparent comparative statics that are useful for interpretation and for empirical design.

First, the education-dominance threshold $\bar{a}^s_{t+1}$ increases with s_0 and ϕ. Higher education costs raise the talent level required to justify the educated technology; higher family-asset productivity raises the outside option of crony continuity, producing the same qualitative effect.

Second, the entry threshold for workers' descendants, $a^{a,s}_{t+1}$, increases with s_t and decreases with expected entrepreneurial profitability. Therefore, policies that reduce the fixed component of entrepreneurial education and startup preparation shift entry toward broader portions of the ability distribution among non-entrepreneur dynasties.

Third, the financing threshold $\hat{a}^s_t$ is increasing in current wages in efficiency units. Intuitively, higher wages increase labor-market opportunities and can reduce entrepreneurial bequests net of educational expenses, tightening feasibility for descendants of marginal entrepreneurial dynasties. This mechanism highlights that wage growth is not unambiguously mobility-enhancing when education is financed by dynasty-specific bequests rather than broad credit access.

Finally, threshold interactions matter more than individual thresholds in isolation. A policy can reduce one threshold while leaving the regime unchanged if another threshold remains binding. This is why equilibrium classification in Propositions 2 and 3 is essential: empirical responses should be expected to be nonlinear and regime dependent.

Figure 1 visualizes the central logic of the model in ability space. The first row isolates descendants of entrepreneurial dynasties. At low ability, inherited ties support crony continuity; at high ability, educated management becomes privately optimal; in the intermediate region, wage work and financing constraints can jointly dominate dynastic continuation. The second row isolates worker dynasties. Even when descendants are talented, entrepreneurial entry requires crossing an additional financing threshold tied to parental resources.

The key interpretation is that talent is filtered by multiple gates. A descendant can fail to enter entrepreneurship either because expected profitability is too low or because educational financing is infeasible. These mechanisms are distinct in theory and in policy. Profitability reforms mainly move occupational cutoffs; educational finance reforms mainly move feasibility cutoffs. The visual map makes clear why equal changes in one threshold can produce very different aggregate responses depending on where the distribution of dynasties is located.

For empirical work, the figure suggests where to look for sharp behavioral margins. Transitions near $a^{a,s}$ and $\bar{\bar{a}}^s$ should be most sensitive to education-cost reforms. Transitions near $\bar{a}^s$ and $a^{\phi,s}$ should be most sensitive to changes in inherited managerial rents or wage conditions. This separation provides a practical guide for identifying which channel dominates in a given institutional context.

For workers' dynasties, bequests are

$$\hat{b}^i_{w,t} = \lambda w_t a^i_t. \tag{23}$$

Hence the parental-ability cutoff for financing entrepreneurial education is

$$\bar{\bar{a}}^s_t = \frac{s_t}{\lambda w_t}, \tag{24}$$

and the wage threshold below which all workers' descendants are constrained is

$$\bar{\bar{w}}^s_t = \frac{s_t}{\lambda}. \tag{25}$$

Proposition 3 (Descendants of workers). *(a) If $w_t < \bar{\bar{w}}^s_t$, all workers' descendants are constrained and become workers.*

(b) If $w_t \geq \bar{\bar{w}}^s_t$, only descendants with $a^i_t \geq \bar{\bar{a}}^s_t$ can finance education; unconstrained descendants choose between entrepreneurship and wage work according to $a^{a,s}_{t+1}$.

Proposition 3 describes the external mobility margin in the model. If wages are below $\bar{\bar{w}}^s_t$, worker dynasties are uniformly excluded from entrepreneurial education and there is no upward dynastic transition. Above that threshold, only descendants of sufficiently able (hence higher-income) workers can finance entry.

This proposition explains why educational access and labor-market conditions interact non-linearly. Growth in wages can open entrepreneurship to worker dynasties, but only if wages cross the financing threshold and if the entrepreneurial return cutoff remains favorable. The model therefore predicts mobility "kinks" rather than smooth responses to income growth.

From a distributional perspective, this mechanism implies that early mobility gains are likely to appear first in the upper segment of worker dynasties. This pattern should not be interpreted as weak responsiveness among poorer households; it is a direct consequence of a financing margin crossed gradually across parental-income groups. Over time, if new entrepreneurial entry raises growth and wages, the same mechanism can broaden access further down the distribution.

4 Steady-State Equilibrium and Growth

4.1 Definition

Definition 1. *A steady-state equilibrium is a tuple*

$$\left(n^*, g^*, w^*, a^{\phi,s}, a^{a,s}, \bar{a}^s, \hat{a}^s, \bar{\bar{a}}^s\right)$$

such that: (i) individual occupational choices are optimal given prices and constraints; (ii) labor and entrepreneurship markets clear; (iii) the distribution of constrained dynasties is stationary; and (iv) technology grows at rate g^* implied by the mass of entrepreneurs using educated management.

Definition 1 is useful because it makes clear that a "steady state" in this model is not a no-change economy. Technology can keep growing at a constant rate, firms can keep entering and exiting, and dynastic transitions can continue. What is stationary is the *distributional structure*: thresholds, constrained shares, occupational masses, and growth rates scale consistently with the technology frontier.

This clarification is useful empirically. An economy can display stable aggregate growth and a stable family-firm share while still experiencing substantial reallocation in underlying dynastic transitions. Flow measures therefore carry more structural content than static shares. Tracking entry, exit, and within-entrepreneur upgrading is essential to distinguish genuine convergence toward an efficient regime from persistence in a low-mobility configuration.

4.2 Flow decomposition of equilibrium dynamics

To make the equilibrium mechanics explicit in the main text, define n_t as the mass of entrepreneurial dynasties at t. The accounting law of motion is

$$n_{t+1} = n_t - X_t + E_t, \tag{26}$$

where X_t is entrepreneurial exit and E_t is entry from worker dynasties.

In the mobility regime, entry comes from unconstrained worker descendants whose ability exceeds $a^{a,s}_{t+1}$:

$$E_t = (1 - n_t)(1 - \bar{\bar{a}}^s_t)(1 - a^{a,s}_{t+1}). \tag{27}$$

Exit comes from entrepreneurial descendants who either prefer wage work or are constrained and cannot adopt educated management:

$$X_t = n_t\left[(1 - a^{\phi,s}_{t+1}) - (\bar{\bar{a}}^s_t - \tilde{a}^s_t)(1 - a^{a,s}_{t+1})\right]. \tag{28}$$

Imposing stationarity in (26) gives the closed-form expression for n^* reported in Proposition 4.

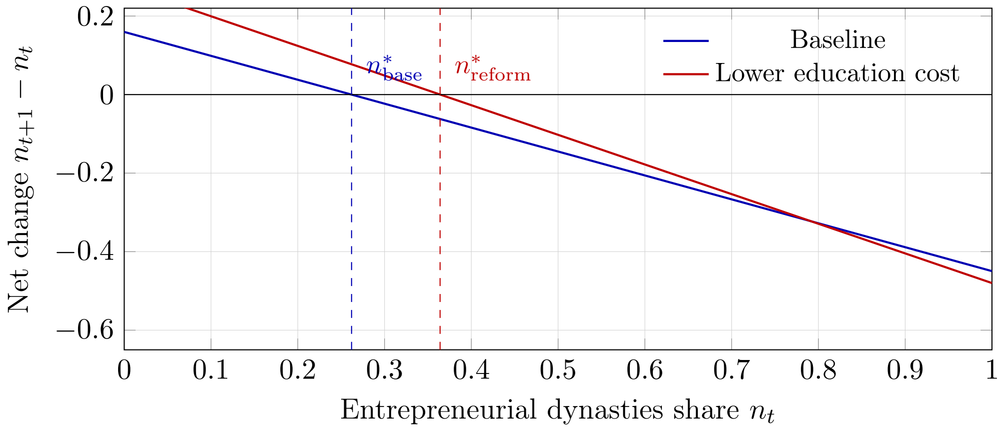


Figure 2: Phase diagram from the flow law. The lines plot $n_{t+1} - n_t$ as a function of n_t under two parameterizations of educational costs. Data are generated from (27)–(28) using a stylized numerical calibration documented in the replication script.

Growth is pinned down by the mass of entrepreneurially educated managers:

$$g_{t+1} = \chi m^a_{t+1}, \qquad m^a_{t+1} = (1 - a^{a,s}_{t+1}) \left[(1 - n_t)(1 - \bar{\bar{a}}^s_t) + n_t(1 - \tilde{a}^s_t)\right], \tag{29}$$

with normalization $\chi = 1$ in the baseline calibration. Hence, "more entrepreneurship" does not automatically imply faster growth: what matters is the composition of entrepreneurs by management type.

Finally, define upward mobility from worker to entrepreneur dynasties:

$$\mathcal{M}_t \equiv \Pr(o_{t+1} = e \mid o_t = \omega) = (1 - \bar{\bar{a}}^s_t)(1 - a^{a,s}_{t+1}). \tag{30}$$

Equation (30) provides a compact empirical mapping: any policy that lowers educational financing thresholds or entrepreneurial entry cutoffs raises mobility one-for-one through this product term.

The same decomposition also clarifies policy diagnostics. Education-finance interventions primarily act through the entry margin E_t by expanding the mass of unconstrained worker descendants. Governance interventions can affect both X_t and the quality composition of remaining entrepreneurs by changing managerial payoffs. Similar trajectories for the aggregate entrepreneurial share can therefore hide very different underlying mechanisms if inflow and outflow are not separately measured.

Figure 2 translates the accounting law of motion into a phase-diagram representation. The zero-crossing of each curve identifies the corresponding steady-state entrepreneurial share. A downward shift in education costs moves the entire curve upward in the relevant range and shifts the fixed point rightward, from approximately $n^* = 0.26$ to $n^* = 0.36$ in this illustrative calibration. The mechanism is transparent: lower costs enlarge the unconstrained pool of potential entrants and reduce exclusion in the worker-to-entrepreneur margin.

The figure is useful because it separates level and speed effects. The slope around the fixed point governs transition speed, while the intercept shift governs long-run level differences. In this model, access reforms can change both objects at once: they alter long-run composition and also steepen short-run adjustments by moving dynasties across binding thresholds. This is why policy evaluation should track dynamic trajectories rather than comparing only pre- and post-reform steady states.

4.3 Polarization equilibrium with social mobility

When family assets are valuable but not dominant and education costs are moderate, the economy displays endogenous firm-type polarization.

Proposition 4 (Polarization equilibrium)**.** *If*

$$\phi < \bar{\bar{\phi}}^s \equiv \left(\frac{s_0\gamma}{\theta(1-\gamma)}\right)^{1/(1-\alpha)} \quad \textit{and} \quad s_0 < \frac{\theta(1-\gamma)}{\gamma^{\,1-\alpha(1-\alpha)}},$$

there exists a steady state with positive entry and exit of firms, positive growth, and positive mobility:

$$n^* = \frac{(1-\bar{\bar{a}}^s)(1-a^{a,s})}{(1-a^{\phi,s})-(\bar{\bar{a}}^s-\tilde{a}^s)(1-a^{a,s})}, \tag{31}$$

$$g^* = \frac{(1-a^{a,s})(1-a^{\phi,s})(1-\bar{\bar{a}}^s)}{(1-a^{\phi,s})-(1-a^{a,s})(\bar{\bar{a}}^s-\tilde{a}^s)}, \tag{32}$$

where $\tilde{a}^s = a^{\phi,s}$ *if* $\hat{a}^s \in [\bar{a}^s, a^{a,s})$ *and* $\tilde{a}^s = \hat{a}^s$ *if* $\hat{a}^s \in [a^{a,s}, 1]$.

Proposition 4 is the core result of the paper. It shows that equilibrium polarization is not an ad hoc assumption, but an endogenous consequence of three threshold systems interacting at once: financing constraints, occupational payoffs, and management-mode choice. The coexistence of well-managed and badly managed family firms appears because dynastic persistence and meritocratic upgrading are both active, but on different subsets of dynasties.

The formula for n^* reveals this directly. The numerator is the potential inflow from worker dynasties that are unconstrained and sufficiently able; the denominator adds the outflow pressure from entrepreneurial dynasties that switch to wage work. Policy and structural parameters affect equilibrium primarily by changing these two extensive margins. This interpretation is useful for empirical work: entry and exit rates contain direct information about latent financing and management frictions.

The proposition says that when inherited ties are productive but not overwhelming, and entrepreneurial education is costly but not prohibitive, the economy settles into a mixed regime. Some family firms upgrade and grow through talent; others survive through inherited rents. Mobility exists, but it is incomplete. This mixed equilibrium is likely to be relevant for economies in intermediate development stages. Organizational forms with very different managerial quality can coexist for long periods without immediate convergence to either full meritocracy or full dynastic closure. In such settings, modest policy shifts can have large medium-run effects if they move a significant mass of dynasties across financing or profitability thresholds.

Proposition 5 (Education costs and growth)**.** *An increase in* s_0 *raises the mass of constrained descendants and lowers:*

(i) the equilibrium number of firms n^**,*

(ii) the share of firms managed through entrepreneurial human capital,

(iii) the long-run growth rate g^*.

Proposition 5 highlights why education policy has macro effects even in a model where technology also depends on incumbent entrepreneurs. Higher s_0 does not only reduce entry from worker dynasties; it also changes internal upgrading within entrepreneurial dynasties by making educated management less attainable. The growth effect is therefore reinforced by both the external and internal margins.

From a policy perspective, this proposition implies that interventions reducing private entrepreneurial education costs should increase growth disproportionately in economies near financing thresholds. Near-threshold economies exhibit large "composition elasticities": small cost changes move many dynasties from constrained to unconstrained states.

An immediate corollary is that policy elasticities are distribution-dependent. Two economies with the same average education cost can respond very differently if one has many dynasties close to the financing margin and the other does not. This is why calibration and empirical evaluation should focus on the distribution of dynastic resources, not only on average policy parameters.

4.4 Low-mobility equilibrium with positive but lower growth

Proposition 6 (Dynastic trap equilibrium)**.** *If*

$$\phi \geq \tilde{\phi}^s \equiv \frac{s_0\gamma}{(1-\gamma)\theta} \quad \text{and} \quad s_0 < \frac{\theta(1-\gamma)}{\gamma^{\,1-\alpha(1-\alpha)}},$$

there exists a steady-state equilibrium with no worker-to-entrepreneur mobility:

$$n^* = n_t, \qquad g^* = n^*(1-\bar{a}^s),$$

which is strictly lower than growth in Proposition 4 for given n^*. All workers' descendants become workers; entrepreneurs' descendants remain entrepreneurs using crony management if $a^i < \bar{a}^s$ and educated management otherwise.

Proposition 6 formalizes a dynastic segmentation trap. The important point is that growth does not collapse to zero; rather, growth becomes *selectively dynastic*. Innovation is carried only by the educated subset of incumbent entrepreneurial lineages, while talented descendants from worker dynasties remain excluded.

This has two implications. First, observing positive growth is not sufficient evidence of efficient talent allocation. Second, persistence in class structure can coexist with moderate aggregate performance for long periods, making the trap politically and empirically difficult to detect without intergenerational mobility data.

A further implication is that the trap can be locally stable even when individual behavior is fully optimizing. The inefficiency does not come from errors by agents; it comes from equilibrium constraints embedded in dynastic finance and inherited control rights. Policy must therefore change feasible sets or payoff structures to move the economy across regimes.

4.5 Welfare and misallocation

To summarize inefficiency, consider a benchmark in which education is universally financeable and occupations depend only on descendants' ability. Let g^{FB} denote the corresponding growth rate and g^* the constrained equilibrium growth rate. The dynamic inefficiency wedge is

$$\Delta_g \equiv g^{FB} - g^* \geq 0,$$

with strict inequality whenever a positive mass of high-ability descendants is excluded from educated entrepreneurship by bequest constraints.

Static inefficiency can be represented by an assignment gap in expected managerial quality:

$$\Delta_m \equiv \mathbb{E}[m^{FB}] - \mathbb{E}[m^*],$$

where m denotes effective managerial input at the firm level. In the model, Δ_m rises with both s_0 and ϕ, but through different channels. Higher s_0 excludes constrained high-ability descendants from the high-productivity technology. Higher ϕ increases the private value of retaining low-ability heirs in control through family assets. The two channels reinforce each other when expensive education coexists with highly productive inherited ties.

4.6 Transition dynamics and persistence

The steady-state characterization also implies transition patterns after policy or structural shocks. Following a permanent decline in s_0, constrained shares fall first among dynasties near the financing margin. Entry by educated entrepreneurs then increases, raising wages and relaxing constraints further in later cohorts. Hence mobility effects are cumulative and may appear with delay.

By contrast, an increase in ϕ can generate persistence. Once low-ability dynasties remain active in larger numbers, the distribution of parental occupations and bequests shifts in a way that keeps mobility low even after ϕ partially reverts. The economy then exhibits history dependence through dynastic composition, not only through contemporaneous incentives.

Figure 3 reports a stylized simulation based directly on the flow and growth equations in Section 4. The experiment compares a baseline path with a reform path featuring a permanent reduction in education costs at period $t = 10$. Two patterns emerge. First, the entrepreneurial share rises gradually rather than discontinuously. This reflects the intergenerational structure of the model: feasibility expands immediately for marginal dynasties, but compositional reallocation propagates cohort by cohort through entry and bequests.

Second, growth responds through both an extensive and an intensive channel. The extensive channel is higher entrepreneurial entry from previously constrained worker dynasties. The intensive channel is a larger share of entrepreneurs operating with educated management, as implied by the decomposition in (29). In the simulation, these channels reinforce each other, producing a persistent growth gap relative to the no-reform path rather than a short-lived spike.

The broader implication is methodological as well as economic. Reduced-form event windows can miss much of the adjustment because transitional dynamics are gradual by construction.

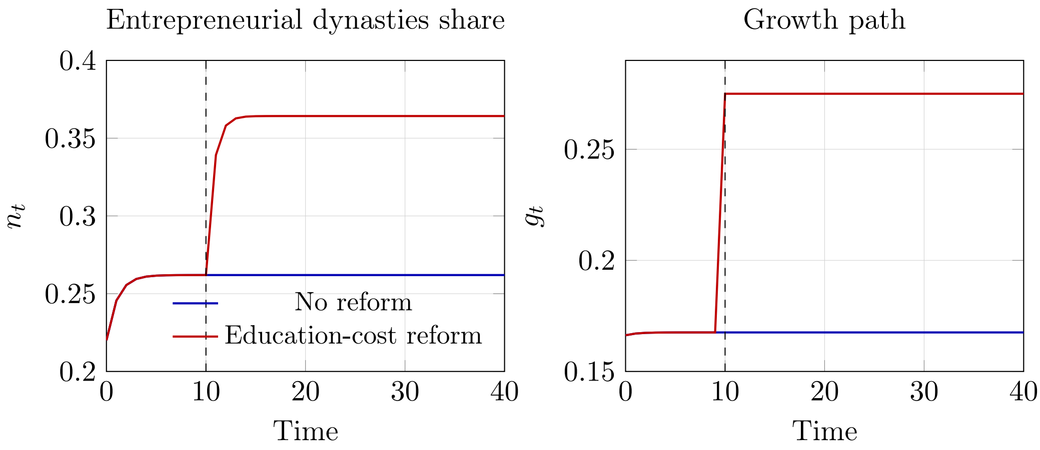


Figure 3: Transition simulation after a permanent reduction in entrepreneurial education costs at $t = 10$ (dashed vertical line). Left panel: path of entrepreneurial-dynasty share. Right panel: implied growth path from (29).

A policy that appears modest in the first few periods can have materially larger effects once intergenerational composition adjusts. This reinforces the paper's central message that social mobility and growth effects must be evaluated jointly and over medium horizons.

4.7 Institutional interpretation across development stages

The model can be read as a reduced-form bridge between institutional quality and dynastic allocation. In low-development environments, formal contract enforcement and managerial labor markets are often thin, raising the relative value of inherited reputation and personalized exchange. In model terms, this corresponds to a higher effective ϕ. At intermediate stages, the same mechanism may support business continuity when formal markets are incomplete. At high development stages, however, a persistently high ϕ can prevent reallocation toward merit-based management and reduce growth.

Education systems and financial institutions shape the second key primitive, s_0. Countries with high private costs of entrepreneurial training and weak educational finance effectively increase s_0, amplifying dynastic persistence. Conversely, broad access to managerial and technical training lowers s_0 and can offset the persistence generated by family-specific rents.

This interpretation implies that similar family-firm prevalence may hide different underlying regimes. A country may have many family firms because inherited assets are productive and education is accessible (high ϕ, low s_0), or because inherited assets are productive and education is inaccessible (high ϕ, high s_0). Only the second case necessarily implies low mobility and larger dynamic inefficiency.

4.8 Policy relevance, empirical agenda, and interpretation

The model implies that family firms are not inherently beneficial or harmful for development; their aggregate role depends on whether inherited family assets complement entrepreneurial investment or substitute for it. In early development phases, inherited relationships can stabilize production when formal institutions and managerial markets are thin. Closer to the frontier, the same mechanism may hinder meritocratic matching if inherited rents dominate educational upgrading.

The formal results also suggest concrete empirical predictions. The share of family firms, taken in isolation, is not a sufficient statistic. What matters is the within-family distribution of management quality and the extent of worker-to-entrepreneur transitions across parental wealth groups. The model predicts stronger growth and mobility responses to policy in economies where financing thresholds bind for a large mass of potentially high-ability descendants.

In policy terms, the framework points to complementarity across reforms. Lower private costs of entrepreneurial education, broader household liquidity, and governance arrangements that raise returns to merit-based succession reinforce one another. Isolated interventions can improve one margin while leaving another margin blocked; combined reforms shift both access and selection.

The framework has limitations that help organize future research. We do not model a corporate control market explicitly, ability is iid across generations, and family-specific assets are represented by a reduced-form parameter. These simplifications keep the mechanism transparent but motivate richer empirical and structural work on the components of inherited managerial rents and on policy designs that jointly affect schooling access and dynastic governance.

5 Economic Interpretation and Policy Design

5.1 How to read the model's core mechanism

The model can be summarized as a dynamic interaction between three state variables: dynastic occupational status, ability realizations, and inherited financial capacity. Ability determines the productivity of entrepreneurial human capital, but inherited financial capacity determines whether this productive option is feasible. Occupational status matters because it influences both available managerial technologies and the bequest process. This three-way interaction generates a persistent but non-degenerate sorting process.

From this perspective, the threshold system is not a technical artifact but a compact representation of structural margins that can be mapped to data. Threshold $\bar{a}^s$ captures internal selection between meritocratic and inheritance-based management among entrepreneurial descendants. Threshold $\hat{a}^s$ captures financing feasibility conditional on parental occupational history. Threshold $\bar{\bar{a}}^s$ captures worker-dynasty access to entrepreneurial tracks. The joint movement of these objects determines how much of the talent distribution is transformed into productive entrepreneurship.

An important implication is that identical ability distributions can produce very different growth trajectories across economies. What differs is the mapping from ability into feasible occupational and managerial choices. In low-access environments, high-ability descendants of worker dynasties can remain excluded for multiple generations, while low-ability entrepreneurial descendants can remain incumbents because family-specific rents preserve dynastic control. The resulting mismatch is dynamic: it affects future bequests and therefore future feasible sets.

5.2 Interpreting threshold nonlinearities for policy

Because occupational sorting is threshold-based, policy effects are inherently nonlinear. A small subsidy to entrepreneurial education can generate limited aggregate effects when most dynasties remain far below the financing margin. The same subsidy can generate large effects when many dynasties are clustered near the threshold. This property implies that policy timing and targeting matter at least as much as average spending intensity.

The model also implies that policy bundles dominate isolated interventions. A pure education-cost intervention lowers s_0 and relaxes financing constraints, but if inherited managerial rents remain very high, low-merit continuity can still absorb a large share of entrepreneurial slots. A pure governance reform that lowers effective ϕ can improve internal selection among incumbents, but without educational access many high-ability outsiders still cannot enter. Joint interventions alter both access and selection, generating multiplicative gains in mobility and growth.

This complementarity has a practical implication for reform sequencing. In environments with severe exclusion of worker dynasties, reforms that first expand educational access may be necessary to create an entrant pool. In environments where access exists but dynastic incumbency is excessively protected, governance reforms may have stronger short-run returns. The model does not prescribe a single universal sequence, but it predicts that long-run gains are largest when both margins are ultimately addressed.

5.3 Welfare accounting and quantitative implementation

The welfare losses implied by constrained sorting can be represented by a static allocation component and a dynamic growth component. A practical reduced-form object is

$$\mathcal{L} = \omega_m \Delta_m + \omega_g \Delta_g,$$

where Δ_m is the managerial-allocation gap and Δ_g is the growth gap relative to a benchmark with broad educational access. Weights ω_m and ω_g can be calibrated from consumption-equivalent welfare metrics or policy-specific discounting criteria.

In quantitative work, the model can be estimated with simulated method of moments by matching the share of family firms, within-family management dispersion, transition rates by parental background, and long-run productivity growth. Counterfactual simulations then evaluate marginal and joint reforms. Particularly informative exercises include the dynamic response of mobility to sustained tuition subsidies, governance reforms that lower returns to low-merit succession, and combined policies that simultaneously alter access and selection.

Transition dynamics are central for policy evaluation. Subsidy-only reforms typically raise entry first and quality later, while governance-only reforms can initially increase exits before high-quality entry adjusts. Joint reforms reduce these transitional trade-offs by broadening the entrant pool while improving the selection criterion applied to control rights. For policy-makers, the model therefore recommends evaluating not only long-run steady states but also medium-run transition paths in mobility and firm composition.

5.4 Distributional implications across dynasties

The model has distributional implications that go beyond occupational status. Because educational access is financed dynastically, shocks to entrepreneurial rents and education costs alter both the level and persistence of wealth inequality. When s_0 rises, constrained dynasties accumulate less entrepreneurial income and therefore leave lower bequests, which tightens future constraints and amplifies persistence. When ϕ rises, low-ability entrepreneurial dynasties can preserve control and maintain above-wage incomes despite weak managerial quality. In both cases, the bequest channel turns short-run sorting distortions into long-run distributional divergence.

This mechanism implies that mobility and inequality should be analyzed jointly. A policy can increase average income while worsening intergenerational mobility if gains accrue mainly to unconstrained dynasties. Conversely, policies that relax financing constraints at the lower end of the dynasty distribution can improve mobility with relatively modest short-run effects on aggregate output, but large medium-run effects through composition. The relevant welfare comparison is therefore dynamic and distribution-sensitive, not purely static.

The framework also predicts asymmetries across the ability distribution. Near the lower tail, changes in s_0 mostly affect feasibility, because descendants are often far from entrepreneurial profitability cutoffs. Near the upper tail, changes in s_0 and ϕ affect both feasibility and managerial mode choice. This generates heterogeneity in policy incidence that standard representative-agent models cannot capture. In empirical work, this suggests focusing on transitions around the upper-middle part of the ability or parental-income distribution, where threshold crossing is most likely.

5.5 Institutional heterogeneity and external validity

Cross-country applicability depends on how institutions shape the primitives of the model. Educational systems and household finance institutions map into effective s_0 by changing private costs and liquidity constraints. Governance quality, contract enforcement, local social capital, and product-market informality map into effective ϕ by changing the private value of inherited relationships. Countries can therefore display similar observed family-firm shares while lying in very different regions of the (s_0, ϕ) space.

This observation matters for external validity of policy advice. A tuition subsidy calibrated in a low-ϕ setting may have limited effects in a high-ϕ setting if governance frictions remain unaddressed. Similarly, governance reforms may underperform in high-s_0 settings where potential entrants remain excluded. The model therefore recommends a diagnostic-first strategy: estimate where the economy sits in threshold space, then design policy packages consistent with that location.

For researchers, the model offers a bridge between macro growth accounting and micro intergenerational data. Macro moments discipline long-run equilibrium objects such as g^* and entrepreneurial shares; micro linked-parent-child data discipline transition equations and financing thresholds. Combining both levels is essential, because either level alone can misclassify regimes. An economy with moderate growth and high family-firm prevalence can either be close to an efficient polarized equilibrium or trapped in dynastic segmentation; without transition and managerial-quality data, these cases are observationally close at the aggregate level.

For policy evaluation, this also implies that success metrics should be multi-dimensional. Relevant indicators include not only growth and firm entry rates, but also within-family management dispersion, worker-to-entrepreneur transition probabilities by parental wealth, and the share of educated managers among entrepreneurial dynasties. Tracking this vector over time is the most direct way to detect whether reforms are changing the underlying sorting mechanism rather than only its short-run aggregates.

5.6 Broader development implications

The model also contributes to a broader debate in development economics on the relationship between social structure and productivity growth. Family connections are often interpreted either as a friction or as an institutional substitute in environments with weak formal markets. The framework developed here reconciles these views by showing that both interpretations can be correct, but in different regions of parameter space and at different stages of development. In low-capacity institutional settings, family assets can sustain firm continuity and reduce transaction risk. As institutions deepen and frontier technologies become more skill-intensive, the same assets can delay reallocation toward higher-productivity managerial matches if educational access remains uneven.

This perspective helps interpret why similar reforms can generate heterogeneous outcomes across countries. Policies that reduce tuition and credit constraints can be highly effective where entry bottlenecks dominate, but less effective where dynastic incumbency rents remain strong. Conversely, governance reforms that alter succession incentives can have rapid effects where talented outsiders already have educational access, but limited effects where access remains tightly linked to parental wealth. The model therefore suggests that policy diagnostics should explicitly separate "access frictions" from "selection frictions" rather than treating family-firm prevalence as a single sufficient statistic.

Finally, the framework offers a concrete interpretation of slow institutional transitions. Even when reforms begin to lower s_0 or ϕ, dynastic distributions adjust gradually through the bequest channel, so observable mobility improvements can appear with a delay. This delay is not a failure of policy design; it is a structural implication of intergenerational state dependence. For this reason, evaluation windows that are too short can systematically underestimate the gains from reforms that target educational access and merit-based succession simultaneously.

6 Extended Discussion of Mechanisms and Transitions

6.1 Regime transition logic in economic terms

The threshold structure implies that regime transitions are not smooth movements along a single margin but coordinated shifts in multiple margins. A decline in education costs, for example, does not only reduce one financing cutoff; it changes the feasible set for worker dynasties, the management mode selected inside entrepreneurial dynasties, and the equilibrium wage path that feeds back into subsequent bequests. The result is a sequence of adjustments in which occupational composition, managerial quality, and growth co-evolve.

This perspective is useful to interpret empirical reform episodes where outcomes appear delayed or nonlinear. In the short run, measured entrepreneurship may rise modestly if new entry is concentrated among dynasties near the feasibility margin. In the medium run, as those entrants accumulate and wages adjust, the same policy can trigger broader reallocation and stronger growth effects. The model therefore predicts that transition elasticity is state-dependent and time-varying rather than constant.

The same reasoning applies in reverse for adverse shocks. If private educational costs rise or household liquidity tightens, the immediate effect can look limited in aggregate data. But once lower entry quality and weaker internal upgrading reduce future bequests, exclusion propagates across cohorts and the economy may drift toward a low-mobility configuration. This propagation mechanism helps explain why temporary disruptions in education finance can have persistent effects on entrepreneurial selection.

6.2 Dynamics of inequality, mobility, and selection

Because education is financed through bequests, inequality and mobility are jointly determined. In the model, dynasties with sustained access to entrepreneurial human capital generate higher expected net income and stronger bequest capacity, which relaxes future constraints for descendants. Dynasties excluded from the same investment channel experience the opposite dynamic. The resulting divergence is not purely a level effect; it is a difference in intergenerational slope.

This channel generates a specific prediction: policies that reduce financing frictions should have larger medium-run effects on mobility than on contemporaneous inequality. Initially, income rankings may remain sticky because incumbent dynasties still benefit from inherited rents. Over time, however, broader access to entrepreneurial education changes who enters and who upgrades managerial practices, and this gradually alters rank transitions across dynasties. For policy evaluation, this implies that transition matrices are often more informative than static inequality moments in early reform years.

An additional implication concerns the top of the ability distribution. High-ability descendants from constrained backgrounds are the most likely to switch from latent entrepreneurial potential to realized wage careers when financing thresholds bind. This reduces the upper-tail quality of entrants and lowers the intensity of talent-based selection in entrepreneurship. In macro terms, growth slows not because ability disappears, but because ability is reallocated toward occupations with lower innovation externalities.

6.3 Family assets, organizational continuity, and misallocation

The parameter ϕ captures a broad set of inherited organizational advantages. In weak-market environments, these advantages can be efficiency-enhancing because they substitute for missing formal institutions. However, the model emphasizes that the same mechanism can become dynamically distortionary when inherited rents are strong enough to sustain low-merit control in settings where skill-intensive management is increasingly valuable.

This dual role helps reconcile contrasting empirical findings on family firms. Studies documenting resilience and continuity are consistent with the low-institutional-capacity region of the model, where family assets support production and reduce contractual frictions. Studies documenting management gaps and slower productivity growth are consistent with regions where inherited rents crowd out meritocratic reallocation. The model does not force a single verdict on family ownership; it provides conditions under which each verdict is likely.

The key empirical object is therefore the interaction between inherited rents and educational access. High ϕ need not be harmful when s_0 is low and talented outsiders can still enter or incumbent dynasties can still upgrade management. High ϕ becomes problematic when combined with high s_0, because both internal and external selection margins weaken simultaneously. This complementarity is central for comparative institutional analysis.

6.4 Policy sequencing and transitional trade-offs

The model suggests that policy sequencing should be tied to the dominant distortion in the initial regime. Where worker-dynasty exclusion is severe, reducing education costs and easing liquidity constraints are likely to be first-order priorities because they create an entrant pool that can compete for entrepreneurial positions. Where access is less constrained but dynastic incumbency is strongly protected, governance and succession reforms may produce larger immediate allocation gains.

Even with appropriate sequencing, transition trade-offs remain. Access reforms may increase the number of potential entrants before governance reforms improve selection among incumbents, generating a temporary phase of broad but noisy entry. Governance reforms may induce faster exit of low-quality incumbents before entry expands, temporarily reducing entrepreneurial mass. The model implies that these transition patterns should be interpreted as adjustment dynamics rather than policy failure.

Combined reforms reduce such trade-offs by acting on both sides of the assignment problem. Access reforms broaden the supply of qualified potential entrepreneurs; governance reforms increase demand for merit-based managerial allocation by reducing rents from low-merit continuity. When implemented together, they accelerate convergence toward a regime with higher mobility and stronger growth. In this sense, complementarities are not only long-run properties but also transition-stabilizing mechanisms.

6.5 Political economy and implementation constraints

Although the model is not explicitly political, its structure implies several implementation frictions that matter for real-world reform design. Dynasties benefiting from high inherited-control rents have incentives to oppose governance reforms that tighten merit-based succession, while dynasties excluded from entrepreneurial education may support access reforms but face weaker organizational power. This asymmetry can generate partial reforms that target only one margin, even when joint reform would be socially more efficient.

The framework helps clarify why such partial reforms often underperform. If policy lowers educational costs without changing succession incentives, more potential entrants are created but allocation of control rights may remain weakly meritocratic. If policy changes governance rules without relaxing educational finance, potential high-ability outsiders can still be excluded from entry. In both cases, measured progress can be positive but smaller than expected because one binding threshold remains in place. This prediction is especially relevant when evaluating gradual reform packages implemented across electoral cycles.

Implementation capacity also affects transition paths. Education-finance reforms require administrative systems that can target constrained households and sustain support over multiple cohorts. Governance reforms require legal enforcement capacity, transparency in succession arrangements, and credible dispute-resolution institutions. The model predicts that weak implementation reduces reform complementarity by making one margin move faster than the other. For policymakers, this implies that institutional capacity is not an external detail but part of the structural environment shaping equilibrium transitions.

Finally, the model suggests that communication and expectations can influence medium-run outcomes even when primitives are unchanged in the short run. If households expect reforms to persist, educational investments and occupational plans may adjust earlier, strengthening dynamic responses. If reforms are perceived as temporary, dynasties may postpone costly transitions, and measured effects remain muted. While this channel is outside the baseline equations, it is consistent with the intergenerational logic of the framework and motivates empirical work on policy credibility and dynastic decision-making.

7 Conclusion

This paper extends Carillo et al. [2019] (*Journal of Economic Growth*) by embedding bequest-financed monetary costs of entrepreneurial education in a dynastic model of family firms. The key implication is a richer selection mechanism in which occupational outcomes depend jointly on descendants' talent and parental financing capacity.

The model delivers two long-run regimes with different development consequences. In the polarization regime, the economy combines endogenous entry and exit, coexistence of well-managed and poorly managed family firms, and positive but incomplete intergenerational mobility. In the dynastic-trap regime, worker dynasties are excluded from entrepreneurial entry, mobility collapses, and growth is lower even though innovation by incumbent entrepreneurial lineages persists.

The policy message is that education access and governance design are complementary. Lower private costs of entrepreneurial education, broader household liquidity, and succession rules that reward managerial merit reinforce one another by widening access and improving internal selection at the same time. The empirical message is equally important: aggregate growth alone is not a sufficient indicator of efficient talent allocation, because positive growth can coexist with persistent dynastic segmentation.

The framework remains intentionally parsimonious. It abstracts from an explicit market for corporate control, treats ability as iid across generations, and captures family-specific rents through a reduced-form parameter. These simplifications are useful for identifying the central mechanism, but they also define a clear agenda for further empirical and structural work on how education finance, family connections, and governance institutions jointly shape social mobility and long-run development.

A Proofs and Supplementary Derivations

This appendix collects the algebraic derivations and proof sketches supporting the main text. Each subsection is named according to the formal statement it supports.

A.1 Proof of Lemma 1

Define the wage $\bar{w}^s_{t+1}$ such that $a^{\phi,s}_{t+1} = 1$:

$$\bar{w}^s_{t+1} = \theta A_{t+1}\phi(1 - g_{t+1}).$$

At this wage, $a^{\phi,s}_{t+1}$ starts at one and declines with w_{t+1}, while $a^{a,s}_{t+1}$ starts below one and increases with w_{t+1}. Since $a^{a,s}_{t+1} = 1$ at $w_{t+1} = \tilde{w}^s_{t+1}$, a unique intersection between the two curves must occur in $(\bar{w}^s_{t+1}, \tilde{w}^s_{t+1}]$ when family-asset productivity is not too large. The restriction $\phi \le \bar{\phi}^s$ is exactly the parameter condition ensuring this ordering.

A.2 Proof of Proposition 1

Given (19), crony-lineage bequests are flat in ability, while educated-lineage bequests increase in parental ability. Threshold $\hat{a}^s_t$ solves $\hat{b}^{i,a}_t = s_t$. The classification follows from comparing $\hat{a}^s_t$ with the ability regions choosing crony versus educated management. If $\hat{a}^s_t \le \bar{a}^s_t$, even the minimum educated lineage leaves enough bequest and constraints are absent. If $\bar{a}^s_t < \hat{a}^s_t \le a^{a,s}_t$, educated lineages are unconstrained while crony lineages are constrained. If $a^{a,s}_t < \hat{a}^s_t \le 1$, a subset of educated lineages also becomes constrained. If $\hat{a}^s_t > 1$, constraints are universal.

A.3 Proof of Lemma 2

From (21)–(22), $\hat{\hat{w}}^s_t \le \tilde{\tilde{w}}^s_t$ is equivalent to

$$\phi(1 - g_t) \le \left(\frac{(1 + g_t)\gamma}{1 + \gamma g_t}\right)^{\alpha}.$$

Under $g_t \ge 0$, the right-hand side is bounded below by γ^α, so $\phi \le \gamma^\alpha$ is a sufficient condition.

A.4 Proof of Proposition 2

The proof follows a partition by current wage regimes. If $w_t \le \hat{\hat{w}}^s_t$, bequests from both crony and educated lineages exceed educational costs, so occupational choices depend only on future income comparisons. If $\hat{\hat{w}}^s_t < w_t \le \tilde{\tilde{w}}^s_t$, crony lineages are constrained while educated lineages are partly constrained depending on parental ability relative to $\hat{a}^s_t$. If $w_t > \tilde{\tilde{w}}^s_t$, entrepreneurial lineages are fully constrained, and descendants choose only between crony continuation and wage work.

A.5 Proof of Proposition 3

For worker lineages, $\hat{b}^i_{w,t} = \lambda w_t a^i_t$. If $w_t < \bar{\bar{w}}^s_t = s_t/\lambda$, then $\hat{b}^i_{w,t} < s_t$ for all $a^i_t \in [0,1]$, so entrepreneurial education is infeasible. If $w_t \ge \bar{\bar{w}}^s_t$, only descendants with parental ability above

$\bar{\bar{a}}_t^s = s_t/(\lambda w_t)$ can finance education.

A.6 Proof Inputs for Proposition 4

The polarization equilibrium combines the flow identity and threshold-based entry and exit mappings. Let

$$n_{t+1} = n_t - X_t + E_t.$$

In the mobility regime:

$$E_t = (1 - n_t)(1 - \bar{\bar{a}}_t^s)(1 - a_{t+1}^{a,s}),$$

$$X_t = n_t \left[(1 - a_{t+1}^{\phi,s}) - (\bar{\bar{a}}_t^s - \tilde{a}_t^s)(1 - a_{t+1}^{a,s})\right].$$

Imposing $n_{t+1} = n_t = n^*$ yields (31). Combining this with the educated-manager mass gives (32).

A.7 Proof Inputs for Proposition 5

Comparative statics use implicit derivatives of key thresholds. In particular, $\partial \bar{a}_{t+1}^s/\partial s_0 > 0$, $\partial a_{t+1}^{a,s}/\partial s_0 > 0$, $\partial \hat{a}_t^s/\partial s_0 > 0$, and $\partial \bar{\bar{a}}_t^s/\partial s_0 > 0$ under the maintained assumptions. These derivative signs imply that higher s_0 increases constrained masses and reduces both entry quality and the educated-manager share, which jointly lowers n^* and g^* in equilibrium.

A.8 Proof Inputs for Proposition 6

In the dynastic-trap regime, worker dynasties satisfy $w_t < \bar{\bar{w}}_t^s$, so entrepreneurial education is infeasible for all worker descendants and entry is shut down. Entrepreneurial activity persists inside incumbent dynasties through internal switching between crony and educated management. Growth remains positive because a positive fraction of incumbent descendants still adopt educated management, but growth is lower than in the polarization regime due to missing outsider selection.

A.9 Supplementary Threshold Algebra

For descendants of entrepreneurs, educated management dominates crony management when

$$\Psi_{t+1}(a_{t+1}^i)^{1/\alpha} - s_t \geq \Psi_{t+1}\left[\phi(1 - g_{t+1})\right]^{1/\alpha}.$$

Rearranging yields (13). For crony management, wage-entrepreneur indifference

$$w_{t+1}a_{t+1}^i = \Psi_{t+1}\left[\phi(1 - g_{t+1})\right]^{1/\alpha}$$

gives (14). For educated management, indifference

$$w_{t+1}a_{t+1}^i = \Psi_{t+1}(a_{t+1}^i)^{1/\alpha} - s_t$$

gives (15). These expressions generate the threshold system used throughout the paper.

References

Alberto Alesina and Paola Giuliano. Culture and institutions. *Journal of Economic Literature*, 53(4):898–944, December 2015. doi: 10.1257/jel.53.4.898. URL http://www.aeaweb.org/articles?id=10.1257/jel.53.4.898.

William J. Baumol. Entrepreneurship: Productive, unproductive, and destructive. *The Journal of Political Economy*, 98(5, Part 1):893–921, 1990.

Morten Bennedsen and N. Foss. Family assets and liabilities in the innovation process. *California Management Review*, 58(1):65–81, 2015.

Morten Bennedsen, Kasper Meisner Nielsen, Francisco Perez-Gonzalez, and Daniel Wolfenzon. Inside the family firm: The role of families in succession decisions and performance. *The Quarterly Journal of Economics*, 122(2):647–691, 2007. ISSN 00335533. URL http://www.jstor.org/stable/25098855.

Marianne Bertrand and Antoinette Schoar. The role of family in family firms. *The Journal of Economic Perspectives*, 88(2):73–96, 2006.

Nicholas Bloom and John Van Reenen. Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122(4):1351–1408, 2007.

Nicholas Bloom and John Van Reenen. Why do management practices differ across firms and countries? *The Journal of Economic Perspectives*, 24(1):203–224, 2010.

Mike Burkart, Fausto Panunzi, and Andrei Shleifer. Family firms. *Journal of Finance*, LVIII (5):2167–2201, 2003.

Maria Rosaria Carillo, Vincenzo Lombardo, and Alberto Zazzaro. The rise and fall of family firms in the process of development. *Journal of Economic Growth*, 24(1):43–78, 2019. doi: 10.1007/s10887-019-09163-5.

Francesco Caselli and Nicola Gennaioli. Dynastic management. *Economic Inquiry*, 51(1):971–996, 2013.

Oded Galor and Omer Moav. Ability-biased technological transition, wage inequality, and economic growth. *The Quarterly Journal of Economics*, 115(2):469–497, 2000.

Oded Galor and Daniel Tsiddon. Technological progress, mobility, and economic growth. *The American Economic Review*, 87(3):363–382, 1997.

Oded Galor and Joseph Zeira. Income distribution and macroeconomics. *Review of Economic Studies*, 60(1):35–52, January 1993.

Ethnic fractionalization and aid-corruption linkages: Evidence from African countries

Maria Rosaria Carillo[a], Valentina Chiariello[a], Rita De Siano[a]*

Abstract

The study investigates whether ethnic fractionalization might play a role in the controversial aid-corruption relationship characterizing most of recipient countries in the world. It may exert pressure on the redistributive process within a given society thus clarifying why aid can be sometimes diverted from more productive uses, causing it to be ineffective or even harmful. Heterogeneity of corruption across countries calls attention to the potential non-linearity in the relationship between corruption and its determinants. An appropriate quantile regression approach for panel data applied to 38 African countries shows that foreign aid exerts a direct beneficial effect on corruption only in least corrupt countries and the reverse in more corrupt countries. Ethnic fractionalization, instead, exerts a direct strong detrimental effect on corruption but contributes to modify the effectiveness of aid campaigns depending on the corruption level, thus raising doubts about adequacy of an equal management of international aid around the world.

Keywords: Corruption, Foreign Aid, Ethnic Fractionalization, Quantile regression.

JEL Classification: C21, D73, F35, Z13.

[a] Department of Business and Economics, University of Naples Parthenope, V. Generale Parisi 13, 80132 Naples, Italy.

* Corresponding author: rita.desiano@uniparthenope.it; https://orcid.org/0000-0002-1242-4025.

Study funded by the Unione europea- Next Generation EU, Missione 4 “Istruzione e Ricerca” Componente C2 CUPI53D23002850006
PRIN 2022 “The role of skill- biased technological change in immiserizing growth, spatial polarization and product quality upgrading” – 2022ZKA4PZ

1. Introduction

Historians and sociologists have long considered ethnic diversity, characterized by the presence of different social groups sharing the same religion, language and culture, as a crucial development issue. By contrast, economists later began to consider fractionalization as a potential factor explaining why some countries perform better than others. In this regard, the most widely shared hypothesis was that a large variety of ethnic groups co-existing within the same country could lead to disagreements, favouring conflict or discrimination which would, in turn, set back the country's very development (Easterly and Levine, 1997).

Interest in this anthropological trait, based on the common history of a population, has returned to the fore with the awareness that ethnic diversity, strengthened also by sizable internal and external migration flows, plays an important role in shaping the political and social contexts responsible for backwardness in several areas of the world (Montalvo and Reyal-Querol, 2005, 2021). Evidence is provided by the economic literature that has so far focused on the pressure which ethnic fractionalization exerts on the redistributive process within a given society. Such influence may leave some groups with fewer opportunities and, at the same time, worsen government performance in terms of wealth and well-being outcomes. The presence of contrasts between different groups seems to be decisive in the political mechanisms and leads to a poor provision of public goods such as infrastructures, education and fair policies, all worth for the individual and overall well-being growth (Adam and Tsarsitalidou, 2019; Alesina and Drazen, 1991; Alesina and Spolaore, 1997; Beach and Jones, 2017; Shleifer and Vishny, 1993; Treisman, 2000). The literature contains several contributions demonstrating that fights armed conflict, aimed at conditioning the distribution of resources, trap countries in a lower level of development (Easterly and Levine, 1997). If the subdivision of the groups is reflected also in political organizations, higher fractionalization can reduce trust among actors (Finseraas and Jakobsson 2012), increase conflict (Selway 2011) and favour rent-seeking activities to grab more resources. Svensson (2000), Hodler (2006) and Hodler and Knight (2012), among others, show that, in the presence of rival ethnic groups, time and resources are diverted from productive to unproductive rent-seeking activities. In this regard, Mauro (1995) revealed that conflict between different groups lowers investments, and La Porta et al. (1999) and Collier (2001) found a consequent reduction in government economic performance. In a society characterized by high ethnic fractionalization, bribery and corruption, lobbying and political campaigning, distorted policies as well as intimidation, or even violence and civil conflict, are the most common activities channelling internal public funds as well as foreign aid towards the most influential ethnic groups.

Corruption therefore represents the main channel through which prevailing groups manage to control available resources. The detrimental relationship between fractionalization and corruption is confirmed in several studies (Amegavi et al. 2021). Shleifer and Vishny (1993), for example, contend that countries which are more ethnically diverse are susceptible to a disorganized form of corruption, and Treisman (2000) finds evidence of a positive relationship between the index of ethnic fractionalization and corruption. Torenvlied and Haarhuis (2008), furthermore, find that in sub-Saharan African countries fractionalization may obstruct anti-corruption reforms. A positive correlation between ethnic fractionalization for African countries is highlighted also by Easterly and Levine (1997) while Glaeser and Saks (2006) and Dincer (2008) find that it holds for the US. More recently, Akbari et al. (2019) argued that marriage practices and kinship structures, by strengthening closed groups of related individuals, encourage favouritism and corruptive behaviours.

La Porta et al. (1999), Treisman (2000) and Alesina et al. (2003), however, prove that by controlling for other variables, such as per capita income, the impact of fractionalization on corruption may lose in terms of robustness, and Serra (2006) and Elbahnasawy and Revier (2012) find no significant effects of fractionalization on corruption.

By contrast, Alesina and La Ferrara (2005) argue that an ethnic mix could even favour an increasing variety in abilities, experiences and cultures that proves productive and enhancing for innovation and creativity. In this regard, the United States may represent a valid example of the coexistence of the two opposite faces driven by an increasing ethnically differentiated society.

Our study fits this literature, detecting the causal effect of ethnic fractionalization on corruption. In particular, by referring to a sample of 38 sub-Saharan African countries observed in the period 1998-2015, we explore whether ethnic fractionalization may represent a conditional factor in determining the effectiveness of foreign aid in reducing recipients' level of corruption. The idea is that corruption may increase with the inflow of foreign aid when recipient countries experience a larger variety of rival ethnic groups. The greater the fractionalization, the stronger the effort made through rent-seeking activities aimed at hoarding resources. The contribution of this study is therefore to provide a new channel through which to explain the controversial results reached in various less developed countries by international aid campaigns. The presence of strong ethnic fractionalization may clarify why aid is sometimes diverted from more productive uses, causing it to be ineffective or even harmful (Hodler and Knight, 2012). If this is the case, what becomes important is not only the size of aid but also the way it is managed in the recipient country.

Furthermore, we are interested in verifying whether this process may change according to the recipients' level of corruption. Our paper thus refers to the strand of the literature, initiated by

Burnside and Dollar (2000)[1], that questions the benefit of aid campaigns given conditional aid effectiveness (Asongu and Nwachukwu 2016; Collier 2008; Collier and Dollar 2004; De la Croix and Delavallade 2014; Knack 2001; Moyo 2009; Svensson 2000; Tavares 2003). Collier (2008) with the 'Bottom billion' and Moyo (2009) with 'Dead aid' clearly show that aid is damaging for recipient countries. Moyo, in particular, asserts that recipient countries are often trapped in a vicious circle of corruption-slower economic growth-poverty and that stopping aid flows would be even better than prolonging them. What happens in such cases is that, similar to the "natural resources curse" (Djankov et al. 2008), huge amounts of resources may generate distorted behaviours such as corruption and rent-seeking activities. The greater the resources the more attempts to grab them weakens local pressures for policymakers' accountability, thus lowering the degree of democracy and favouring the establishment of autocratic governments (Bräutigam and Knack 2004). Political elites, indeed, have little incentive to give up exceptional resources ensuring fringe benefits (vehicles, study tours, salary increments, etc.), particularly in low-income countries where additional resources would otherwise not be available.

Our analysis focuses on sub-Saharan African countries just because, as stated by the United Nations Economic Commission for Africa, most frequently in these countries every year a substantial amount of resources appears to be lost due to corruption. Along with this primacy, in 2019 African countries registered the lowest perception of corruption, with a score of 32 out of 100, far below the international average of 43[2]. Hence such a conditional effect evokes serious concerns regarding the potential impact of foreign aid on the socio-economic development of recipient regions. Surveys available for African countries (Public Expenditure Tracking Survey, PETS) reveal the presence of large diversions of grants from their purposes due to governors' corrupt behaviours. Reinikka and Svensson (2005), for example, looking at primary school projects in Uganda, demonstrate that only 13% of aid targeting primary schools was actually employed in the education sector. Asongu and Nwachukwu (2016), looking at the World Governance Indicators (WGI), evidence that aid flows are detrimental both to the control of corruption and the rule of law in recipient countries.

Some scholars demonstrate the presence of endogeneity in the linkage between aid and corruption due to a problem of reverse causality (Knack 2001; Svensson 2000; Tavares 2003). De la Croix and Delavallade (2014), among others, impute this endogeneity to the quality of the institutions and the productivity level. In particular, they suggest that, since heterogeneity in productivity appears to be more important than differences in institutional quality, it would be optimal for more corrupt countries

[1] Other contributions to this strand of literature are those by Hansen and Tarp (2000, 2001), Dalgaard et al. (2004), Angeles and Neanidis (2009), and Doucouliagos and Paldam (2010).
[2] Data are taken form Trasparency International 2020 report.

to receive more foreign aid in order to gain in terms of productivity and hence also in terms of lower corruption. In our investigation, such an endogeneity problem is detected by implementing an IV approach and choosing two different instruments for foreign aid in accordance with the reference literature: infant mortality, as a measure of the recipient's development level, and recipient asylum seekers in donor countries, to account for the motivation of sending countries.

Finally, on the basis of previous studies showing that sub-Saharan African countries are characterized by heterogeneous corruption levels (Amegavi et al., 2021, among others), we detect potential non-linearity in the relationship employing the quantile regression estimator approach for panel data (QRPD) in a framework of instrumental variables. This procedure, developed by Powell (2016), may help us to provide empirical insight into the distributional heterogeneity of the relationship between corruption and its main determinants which, if confirmed, would imply that the "one size fits all" contrast measure could be sometimes inadequate in mitigating corruption.

In particular, we expect the pre-existing distribution of corruption level in recipient countries to influence not only aid direct effectiveness in reducing corruption but also the role played by ethnic fractionalization in conditioning foreign aid capacity to boost recipients' good policies aimed at controlling the corruption level (Billger and Goel 2009; Okada and Samreth 2012). The underlying hypothesis is that an increase in ethnic fractionalization will increase the fight for resources between rival groups, with a marginally higher detrimental impact on the corruption level when the latter is particularly low. By contrast, when corruption is already very widespread an increase in ethnic diversity would not be so detrimental due to an already high pre-existing conflicting context within society. People would not perceive the greater effort of ethnic leaders to grab available resources which is so crucial in worsening the quality of institutions, the population being already disenchanted with the conduct of their rulers. Indeed, in countries showing high levels of corruption, policymaker credibility appears to be notably low with people having very little trust in their fairness.

The main policy implication would be that as countries may be characterized by different levels of corruption, which may also change over time, measures aimed at managing foreign aid to favour growth processes should account for such differences.

Our results show that aid may have a beneficial effect on the perceived corruption level only when recipient countries are characterized by a lower level of pre-existing corruption. Second, ethnic fractionalization increases corruption, which supports previous studies by Shleifer and Vishny (1993) and Treisman (2000). However, the presence of higher ethnic fractionalization makes the use of resources less effective as it may contribute to increase the probability of their misappropriation. The analysis proves the presence of a significant heterogeneity along the distribution of corruption. For lower levels of corruption (up to the 25^{th} percentile) an increasing fractionalization could potentially exacerbate corruption levels because of rent-seeking activities aimed at grabbing available resources, including foreign aid. By contrast, when corruption is more widespread, ethnic diversity may succeed in curbing the perverse effect of aid. On the basis of our results, we may conclude that when corruption is low it would be preferable to have greater polarization and hence higher centralization of decisions regarding the use of international aid aimed at favouring the economic growth of recipients. In the presence of higher levels of corruption, instead, aid effectiveness could increase with ethnic diversity.

Our paper is structured as follows: Section 2 describes the methodology and data, Section 3 presents the results, and Section 4 concludes.

2. Methodology and data

Our analysis aims to detect the impact of foreign aid on corruption and to investigate the role played by ethnic fractionalization in the aid-corruption relationship in recipient countries. The empirical analysis employs a cross-section panel data of 38 sub-Saharan African countries from 1998 to 2015. The values for corruption are obtained from the Worldwide Governance Indicator "Control of Corruption" issued by the World Bank. We rescale this indicator to make the interpretation of the estimated coefficients more intuitive [3]. The main independent variables are foreign aid (*Aid*) and ethnic fractionalization (*EF*). Aid is taken from the OECD Creditor Reporting System and is expressed in terms of the share of recipient countries' GDPs. The degree of ethnic fractionalization, taken from the historical index of ethnic fractionalization dataset by Drazanova (2019)[4], indicates the percentage of ethnic groups in the country and corresponds to the probability that two individuals, taken randomly from the resident population, are not from the same ethnic group. The employment of data taken from Drazanova's dataset allows us to include yearly values of fractionalization for each country, whereas most previous studies include only the value corresponding to a single reference year. The ethnic fractionalization index in each year is obtained from a decreasing transformation of the Herfindahl concentration index measured by:

$$EF_c = 1 - \sum_{i=1}^{n} {S_i}^2$$

[3] Values range from -2.5 to 2.5 and a larger value indicates a higher level of corruption.
[4] The original data are taken from the Composition of Religious and Ethnic Groups (CREG) project initiated by the Cline Center for Democracy, University of Illinois at Urbana-Champaign. The original dataset refers to 165 countries, providing annual values over the period 1945 – 2013.

where EF_c is the level of ethnic fractionalization in country c, i refers to ethnic groups and S_i is the proportion of the population in unit c belonging to ethnic group i ($i=1,...,n$). The ethnic fractionalization index ranges from 0, when there is no ethnic fractionalization and all individuals belong to the same ethnic group, to 1, where each individual belongs to a different ethnic group.

We include the interaction term *Aid*EF* that allows the direct effect of ethnic fractionalization on corruption to be disentangled from the indirect effect due to "good management" of foreign aid captured by some standard covariates accounting for institutional, economic and social factors influencing the level of corruption (Tavares 2003). For the quality of institutions, we include: an index of the degree of democracy (*polity2*), capturing significant national political changes taken from the polity IV project dataset (Marshall 2016); an index of perceptions of political stability (*pol_stab*) by the Worldwide Governance Indicator, which accounts for risks of riots and terrorist incidents. Per capita GDP[5] is included to account for the influence of economic development on corruption (*lnGDP*). In order to encompass short-term business cycle noises and correlation effects, we take three-year timespans for each variable. Table 1 presents the summary statistics for all variables in use.

Table 1. Summary statistics

Variables	Mean	SD	Min	Max
Control of corruption index	.5968333	.5804146	-.98	1.73
Aid	.8238982	.9470751	.0020893	9.358.246
EF	.6743567	.2012668	.0813333	.889
Aid*EF	.5390463	.7759453	.0017822	8.319.481
lnGDP	682.625	.9978974	5.01	9.35
polity2	1.580.583	5.126.506	-9	10
pol_stab	-.5637917	.9189503	-2.83	1.08
british legal origin	.4	.4909218	0	1

All variables are lagged by one period to control for simultaneity problems while time and country fixed effects account for the variability not adequately controlled by the covariates and the omitted variables problem. Finally, robust standard errors are used to treat the heteroskedasticity. The model estimated by means of the Ordinary Least Squares (OLS) procedure is as follows:

$$Corruption_{it} = \beta_0 Aid_{i,t-1} + \beta_1 EFindex_{i,t-1} + \beta_2 (Aid * EFindex)_{i,t-1} + \beta_3 X_{i,t-1} + \theta_i + T_t + \mu_{it} \quad (1)$$

[5] Data on GDP are in constant 2010 US dollars and are taken from the World Bank.

with X indicating the vector of covariates, and T_t, θ_i and μ_{it}, respectively, the time and country fixed effects and the stochastic error term with zero mean and constant variance.

Besides the OLS regression, to account for a potential reverse causality existing between aid and corruption (Alesina and Dollar 2000), we run an IV regression to eliminate other possible sources of endogeneity. As an instrument for aid, we use the infant mortality rate (Mishra and Newhouse 2009) that accounts for the general health conditions usually considered one of the main attractors for foreign aid aimed at ensuring the population's basic needs. In addition, we build a new instrument correlated with donors' internal conditions, to account for social and political pressures behind the donor's decision to send aid. In particular, for each recipient, we select the five major donors and calculate the ratio between their inflows of asylum seekers and domestic population. The average of these ratios represents the second instrument for aid used in the IV regression model. The idea is that higher inflows of refugees increase the host population's awareness of economic problems and humanitarian crises plaguing less developed countries, thus raising donors' willingness to send aid.

To account for the presence of non-linear effects, neglected by the use of standard regression techniques (Bitler et al. 2006), we complete our investigation following the quantile regression methodology. This technique would provide a better understanding of the nuanced relationship between the corruption and its determinants as this approach tests the relationship at any point, conditional on the distribution of the dependent variable. Specifically, this procedure enables us to evaluate the impact of fractionalization, aid and the role played by fractionalization on the influence that aid may exert on corruption, throughout the conditional distribution of corruption, given the fact that it may vary across countries. The quantile regression technique allows us to explain the impact of variation within groups of countries sharing the same level of corruption, thus justifying the presence of controversial effects of fractionalization and aid emerging in previous studies.

We follow the quantile regression panel data approach (QRPD) with non-additive country fixed effects developed for an instrumental variable framework by Powell (2016)[6].The QRPD estimator provides point estimates which can be interpreted in the same manner as cross-sectional regression estimates (i.e., the impact of the explanatory variables on each quantile of the outcome distribution), while allowing an arbitrary correlation between the fixed effects and the instruments, and provides consistent estimates also for small time intervals. Our panel quantile specification is as follows:

$$Q_\tau(Corruption_{it}) = \beta_{0\tau}Aid_{i,t-1}(\mu_{it}) + \beta_{1\tau}EFindex_{i,t-1}(\mu_{it}) + \beta_{2\tau}(Aid * EFindex)_{i,t-1}(\mu_{it}) + \beta_{3\tau}X_{i,t-1}(\mu_{it}) \quad (2)$$

[6] An advantage of using non-additive fixed effects is that, unlike additive fixed effects, it does not alter the distribution of the dependent variable as parameters are assumed to depend only on the error term time-varying components. Rather, it uses within-individual variation for identification purposes and maintains the non-separable disturbance property which typically motivates the use of quantile estimators

where Q_τ is the conditional quantile of corruption such that $\tau = 0.10, 0.25, 0.50, 0.75\ 0.90$, and $\mu_{it}=f(\theta_i, \varepsilon_{it})$ are individual country heterogeneous effects inseparable from the explanatory variables.

First, we estimate models excluding the interest variables accounting for the impact of the ethnic fractionalization and then we repeat the estimation including them. The reason being to verify whether the β-coefficient for aid might be affected by the introduction of these variables and to test for the robustness of the aid-corruption linkage.

3 Results

The panel regression estimates presented in Tables 2 and 3 show the results obtained with the OLS, the IV and the QRPD for the 10th, 25th, 50th, 75th, and 90th percentiles of the conditional corruption distribution. The lower quantiles (10th and 25th percentiles) refer to the relatively least corrupt countries in the time interval. The higher quantiles (75th and 90th percentiles) refer to the most corrupt countries.

Table 2. Estimation results baseline model without fractionalization (eq. 1)

VARIABLES	OLS	IVfe	Q10	Q25	Q50	Q75	Q90
Aid	-0.0816***	-0.0787	-0.235***	0.154**	0.0806	0.0163	-0.0373
	(0.0160)	(0.0722)	(0.0582)	(0.0745)	(0.296)	(0.0123)	(0.170)
lnGDP	-0.475***	-0.474***	-0.534***	-0.216***	-0.679	-0.0805***	-0.674
	(0.159)	(0.123)	(0.0911)	(0.0422)	(0.431)	(0.0109)	(0.424)
polity2	-0.00714	-0.00733	-0.00766	-0.0158*	0.00283	-0.0255***	0.0216
	(0.00724)	(0.00845)	(0.00916)	(0.00851)	(0.0357)	(0.00208)	(0.0241)
pol_stab	-0.0954	-0.0960**	-0.237***	-0.111***	-0.0305	-0.172***	-0.148**
	(0.0753)	(0.0425)	(0.0896)	(0.0315)	(0.0625)	(0.0229)	(0.0707)
Time fixed effects	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Observations	228	228	228	228	228	228	228
R-squared	0.211	0.211					
Number of groups	38	38	38	38	38	38	38

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

The direct impact of aid on corruption, controlling only for institutional, economic and social condition characterizing recipient countries, is indicated in Table 2. OLS regression reveals the presence of a very small reduction in corruption brought about by foreign aid while the IV estimator referring to the whole distribution does not evidence any statistically significant effect of aid. The

QRPD estimator, instead, reveals that at the lowest level of corruption (10^{th} percentile) aid reduces corruption while a positive effect is revealed in the second quantile 25^{th} percentile), after which aid loses its statistical significance. All the covariates included in the model show the expected sign. Among them, development level and political stability of the country show a higher effectiveness in reducing corruption at the extremes of the distribution function. Comparison of the OLS and quantile regression results thus reveals that the impacts of the influencing factors on corruption are heterogeneous. Hence, the effects of foreign aid and the other covariates vary under different levels of corruption.

Results change considerably when introducing the other two variables of interest, namely the index of ethnic fractionalization and the interaction between ethnic fractionalization and foreign aid, as shown in Table 3. The impact of foreign aid in all regressions appears to be higher in comparison to previous model specification, suggesting that, when accounting for the presence of different ethnic groups, aid exerts a stronger impact on the perception of corruption in recipient countries. There is also clearer evidence of a non-linear relationship between aid and corruption as aid estimates are strongly significant and negative in least corrupt countries (10^{th} and 25^{th} percentiles) while it reinforces corruption in most corrupt countries (75^{th} and 90^{th} percentiles).

Ethnic fractionalization (EF) has a significant and positive effect on corruption across the entire corruption distribution, which implies a cumulative effect on corruption that exacerbates corruption levels in sub-Saharan African countries, thus supporting the evidence of previous studies (Shleifer and Vishny, 1993; Treisman, 2000). However, the most detrimental effects are those corresponding to the 50^{th} and 75^{th} percentiles (1.260*** and 1.437***, respectively).

Our main result is perhaps the fact that the beneficial influence of aid in reducing corruption may vanish in the presence of ethnic fractionalization. A highly detrimental impact is found both in the IV estimation, run on the whole distribution, and in the lowest quantiles (10^{th} and 25^{th} percentiles) when using the QRPD. The size of these coefficients almost implies that aid effectiveness could be completely offset by the struggle for the sharing of resources when different ethnic groups co-exist in recipient countries. By contrast, in most corrupt countries the presence of several rival groups ensures a more transparent management of funds, including those collected through aid campaigns.

This outcome is robust also when the model is enriched by accounting for the British legal origin (*british*) of recipient countries[7] that is almost perfectly correlated with relevant background variables

[7] Results are presented in the appendix.

detecting the colonial influence on countries' economic outcomes (Glaeser and Shleifer, 2002; Beck, Demirgüç-Kun and Levine, 2003; Klerman et al., 2011).

All other covariates, when statistically significant, show the expected results. Both economic and social controls help the reduction of corruption in recipient countries, particularly when the pre-existing level of corruption is relatively high.

Table 3. Estimation results including ethnic fractionalization and the interaction term (eq. 2)

VARIABLES	OLS	IVfe	Q10	Q25	Q50	Q75	Q90
Aid	-0.168	-1.174**	-0.432**	-0.162***	0.0412	0.385***	0.139***
	(0.186)	(0.464)	(0.172)	(0.0233)	(0.140)	(0.0159)	(0.0410)
EF	0.921	1.891	0.740***	0.814***	1.260***	1.437***	0.656***
	(1.864)	(1.181)	(0.209)	(0.0885)	(0.180)	(0.0285)	(0.100)
*Aid*EFindex*	0.104	1.194**	0.399**	0.143***	-0.182	-0.467***	-0.201***
	(0.209)	(0.500)	(0.175)	(0.0353)	(0.157)	(0.0190)	(0.0437)
lnGDP	-0.504***	-0.713***	-0.561***	-0.195***	-0.233***	-0.109***	-0.112***
	(0.175)	(0.166)	(0.0911)	(0.00575)	(0.0208)	(0.00571)	(0.0155)
polity2	-0.00335	0.0173	-0.0394***	-0.0265***	-0.0176**	-0.0349***	-0.0457***
	(0.00687)	(0.0124)	(0.00739)	(0.00218)	(0.00740)	(0.000785)	(0.00256)
pol_stab	-0.0521	0.00431	-0.0530	-0.130***	-0.218***	-0.161***	-0.152***
	(0.0726)	(0.0463)	(0.0847)	(0.0105)	(0.0157)	(0.0151)	(0.0161)
Time fixed effects	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Observations	228	228	228	228	228	228	228
R-squared	0.185	-0.255					
Number of groups	38	38	38	38	38	38	38

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In summary, our results suggest that aid itself may be beneficial in reducing the level of corruption, above all at lower levels of corruption, thereby succeeding in abating one of the main factors holding less developed countries in a poverty-trap condition. Such conditions may be improved by appropriate use of international aid inflows aimed at supporting recipients' growth and development. Despite these positive effects, our analysis reveals that if the recipient country is also characterized by strong ethnic differentiation, inflows of financial resources could fuel or exacerbate the conflict between groups, which would worsen the pre-existing state of corruption in less corrupt countries while the opposite would occur in more corrupt countries.

4. Conclusions

Using panel data covering 38 sub-Saharan African countries over the period 1998 to 2015, we sought to ascertain whether ethnic fractionalization could play a role in determining the effectiveness of foreign aid in reducing recipient countries' level of corruption. Our findings reveal that foreign aid does contribute to reduce the level of perceived corruption in recipient countries. However, the direct contribution of aid appears to be smaller in ethnically homogeneous countries while it increases considerably when accounting for ethnic fractionalization. Ethnic diversity exerts a positive effect on corruption across the entire corruption distribution, implying that by favouring more inefficient redistributive policies it contributes to exacerbate the overall corruption level. This is confirmed particularly for least corrupt countries by the positive and highly significant effect associated to the interaction between aid and ethnic fractionalization. This would explain how intensification of conflict between different ethnic groups can adversely affect both the quality of the political process, economic performance and hence the outcome of foreign aid distribution processes. However, when corruption is particularly high, the detrimental effect of fractionalization disappears as it is wholly negligible compared to the low trust and high conflict characterizing social relations. On the contrary, when corruption itself is already high, ethnic differences may contribute to avoid centralization of decisions that would lead to a waste of resources.

In line with the strand of literature focusing on the conditional effectiveness of aid, we provided information on the conditions under which aid is effective in reducing the corruption level. While accounting for a key factor such as ethnic fractionalization would help design effective aid campaigns, it is even more important to bear in mind that development strategies and anti-corruption measures must be based on the level of corruption already existing in the country. The distributional heterogeneity identified in the relationship between corruption and its main determinants suggests that the "one size fits all" measure can at times be inadequate in mitigating corruption. This policy implication may thus be useful to address the many differences in aid effectiveness worldwide.

References

Adam, Antonis and Tsarsitalidou, Sofia (2019) Environmental policy efficiency: measurement and determinants. Economics of Governance 20(1), 1-22.

Akbari, Mahsa, Bahrami-Rad, Duman and Kimbrough, Erik O. (2019) Kinship, fractionalization and corruption. Journal of Economic Behavior & Organization 166, 493-528.

Alesina, Alberto and Dollar, David (2000) Who gives foreign aid to whom and why?. Journal of Economic growth 5(1), 33-63.

Alesina, Alberto and Drazen, Allan (1991) Why are stabilizations delayed?. The American Economic Review 8(1), 1170-1188.

Alesina, Alberto and La Ferrara, Eliana (2005) Ethnic diversity and economic performance. Journal of Economic Literature, 43, 762–800.

Alesina, Alberto and Spolaore, Enrico (1997) On the number and size of nations. The Quarterly Journal of Economics 112(4), 1027–1056.

Alesina, Alberto, Devleeschauwer, Arnaud, Easterly, William, Kurlat, Sergio and Wacziarg Romain (2003) Fractionalization. Journal of Economic growth 8(2), 155-194.

Amegavi, George B., Quarshie, Abigail and Mensah, James K. (2021) Mitigating Corruption in Sub-Saharan Africa: Does Heterogeneity in Corruption Levels Matter?. Public Integrity, 1-14.

Angeles, Luis and Neanidis, Kyriakos C. (2009) Aid Effectiveness: The Role of the Local Elite. Journal of Development Economics 90(1), 120–34.

Asongu, Simplice A. and Nwachukwu, Jacinta C. (2016) Foreign aid and governance in Africa. International Review of Applied Economics 30(1), 69-88.

Beach, Brian and Jones, Daniel B. (2017). Gridlock: Ethnic diversity in government and the provision of public goods. American Economic Journal: Economic Policy, 9(1), 112-36.

Beck, Thorsten, Demirgüç-Kunt, Asli and Levine, Ross (2003). Law and finance: why does legal origin matter?. Journal of comparative economics, 31(4), 653-675.

Billger, Sherrilyn M.and. Goel, Rajeev K. (2009) Do existing corruption levels matter in controlling corruption?: Cross-country quantile regression estimates. Journal of Development Economics 90(2), 299-305.

Bitler, Marianne P., Gelbach, Jonah B Hoynes, Hilary W. (2006) What mean impacts miss: Distributional effects of welfare reform experiments. The American Economic Review 96(4), 988-1012.

Bräutigam, Deborah A. and Knack, Sthephen (2004) Foreign aid, institutions, and governance in sub-Saharan Africa. Economic development and cultural change 52(2), 255-285.
Burnside, Craig and Dollar David (2000) Aid, Policies, and Growth. American Economic Review 90(4), 847–68.
Collier Paul and Dollar David (2004) Development Effectiveness: What Have We Learnt?. The Economic Journal 114(496), 244-271.
Collier, Paul (2001). Ethnic diversity – An economic analysis. Economic Policy, 128–166.
Collier, Paul (2008) The politics of hunger: How illusion and greed fan the food crisis. Foreign Affairs, 67-79.
Dalgaard, Carl-Johan, Hansen, Henrik and Tarp Finn (2004) On the Empirics of Foreign Aid and Growth. Economic Journal, 114(496), 191–216.
De la Croix, David and Delavallade, Clara (2014) Why corrupt governments may receive more foreign aid. Oxford Economic Papers 66(1), 51-66.
Dincer, Oguzhan C. (2008) Ethnic and religious diversity and corruption. Economics Letters: 99(1), 98-102.
Djankov. Simeon, Montalvo, Jose G. and Reynal-Querol, Marta (2008) The curse of aid. Journal of Economic growth 13(3), 169-194.
Doucouliagos. Hristos and Paldam Martin (2010) Conditional Aid Effectiveness: A Meta-study. Journal of International Development, 22(4), 391–410.
Drazanova, Lenka (2019) Historical index of ethnic fractionalisation dataset.
Easterly, William and Levine Ross (1997) Africa's growth tragedy: policies and ethnic divisions. The Quarterly Journal of Economics 112(4), 1203-1250.
Elbahnasawy, Nasr G. and Revier, Charles F. (2012) The determinants of corruption: Cross-country-panel-data analysis. The Developing Economies 50(4), 311-333.
Finseraas, Henning and Jakobsson Niklas (2012) Trust and ethnic fractionalization: the importance of religion as a cross-cutting dimension. Kyklos 65(3), 327–339
Glaeser, Edward L. and Saks, Raven E. (2006) Corruption in America. Journal of Public Economics 90(6-7), 1053-1072.
Glaeser, Edward L. and Shleifer, Andrei (2002). Legal origins. The Quarterly Journal of Economics, 117(4), 1193-1229.
Hansen, Henrik and Tarp, Finn (2000) Aid Effectiveness Disputed. Journal of International Development 12(3), 375–98.
Hodler, Roland (2006) The Curse of Natural Resources in Fractionalized Countries. European Economic Review 50(6), 1367–86.

Hodler, Roland and Knight, David S. (2012) Ethnic fractionalisation and aid effectiveness. Journal of African Economies 21(1), 65-93.

Klerman, Daniel M., Mahoney, Paul G., Spamann, Holger and Weinstein, Mark I. (2011) Legal origin or colonial history?. Journal of Legal Analysis, 3(2), 379-409.

Knack, Sthephen (2001) Aid dependence and the quality of governance: cross-country empirical tests. Southern Economic Journal, 310-329.

La Porta, Raphael, Lopez-de-Silanes, Florencio, Shleifer, Andrei and Vishny, Robert (1999) The quality of government. The Journal of Law, Economics, and Organization 15(1), 222-279.

Marshall, Monty G. (2016) COUP D'ÉTAT EVENTS, 1946-2015 CODEBOOK Marshall Monty G. and Marshall Donna R. Center for Systemic Peace.

Mauro, Paolo (1995) Corruption and growth. The Quarterly Journal of Economics 110(3), 681-712.

Mishra, Prachi, Newhouse, David (2009) Does health aid matter?. Journal of health economics 28(4), 855-872.

Montalvo, Jose G. and Reynal-Querol, Marta (2005). Ethnic diversity and economic development. Journal of Development economics, 76(2), 293-323.

Montalvo, Jose G. and Reynal-Querol, Marta (2021). Ethnic diversity and growth: Revisiting the evidence. Review of Economics and Statistics, 103(3), 521-532.

Moyo, Dambisa (2009) Dead aid: Why aid is not working and how there is a better way for Africa Macmillan.

Okada, Keisuke and Samreth, Sovannroeun (2012) The effect of foreign aid on corruption: A quantile regression approach. Economics Letters 115, 240-243.

Powell, David (2016) Quantile regression with nonadditive fixed effects. Quantile Treatment Effects, 1–28.

Reinikka, Ritva and Svensson Jakob (2005) Fighting corruption to improve schooling: evidence from a newspaper campaign in Uganda. Journal of the European economic association 3(2-3), 259-267.

Selway, Joel S. (2011) Cross-cuttingness, cleavage structures and civil war onset. British Journal of Political Science 41(1), 111-138.

Serra, Danila (2006) Empirical determinants of corruption: A sensitivity analysis. Public Choice: 126(1), 225-256.

Shleifer, Andrei and Vishny, Robert W. (1993) Corruption. The Quarterly Journal of Economics 108(3), 599-617.

Svensson, Jakob (2000) Foreign aid and rent-seeking. Journal of International Economics 51, 437-461.

Tavares, Jose (2003) Does foreign aid corrupt?. Economics Letters 79(1), 99-106.

Torenvlied, René and Klein Haarhuis, Carolien M. (2008). Polarization and policy reform: Anti-corruption policymaking in sub-Saharan Africa. Journal of Peace Research, 45(2), 223-240.
Treisman, Daniel (2000) The causes of corruption: a cross-national study. Journal of public economics 76(3), 399-457.

Appendix

Table 3A. Estimation results when including British legal origin control

VARIABLES	OLS	IVfe	Q10	Q25	Q50	Q75	Q90
Aid	-0.168	-1.174**	-0.991***	-0.175***	0.286***	0.126***	0.131***
	(0.186)	(0.464)	(0.234)	(0.0663)	(0.0450)	(0.0477)	(0.0402)
EF	0.921	1.891	0.445**	0.532***	1.535***	1.342***	0.739***
	(1.864)	(1.181)	(0.226)	(0.0622)	(0.120)	(0.0354)	(0.0406)
*Aid*EF*	0.104	1.194**	1.046***	0.171**	-0.463***	-0.186***	-0.186***
	(0.209)	(0.500)	(0.237)	(0.0781)	(0.0648)	(0.0493)	(0.0353)
lnGDP	-0.504***	-0.713***	-0.478***	-0.219***	-0.195***	-0.185***	-0.131***
	(0.175)	(0.166)	(0.0750)	(0.0152)	(0.0252)	(0.00444)	(0.0159)
polity2	-0.00335	0.0173	-0.0109***	-0.0343***	-0.0365***	-0.0346***	-0.0480***
	(0.00687)	(0.0124)	(0.00378)	(0.00257)	(0.00163)	(0.00170)	(0.00238)
pol_stab	-0.0521	0.00431	-0.0723	-0.0863***	-0.160***	-0.101***	-0.175***
	(0.0726)	(0.0463)	(0.0507)	(0.0308)	(0.0223)	(0.00665)	(0.0235)
british			-0.0535*	-0.167***	-0.192***	0.0124	-0.0387**
			(0.0274)	(0.0258)	(0.0346)	(0.0173)	(0.0159)
Time fixed effects	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Observations	228	228	228	228	228	228	228
R-squared	0.185	-0.255					
Number of groups	3838		38	38	38	38	38

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Financial crisis and the convergence of European welfare provision

Bonasia Mariangela[a], De Siano Rita[b]

[a] University of Naples "Parthenope", Department of Business and Economics, Via G. Parisi 13, Naples (Italy), CNR-ISSM, CRISEI. mariangela.bonasia@uniparthenope.it

[b] University of Naples "Parthenope", Department of Business and Economics, Via G. Parisi 13, Naples (Italy), CRISEI. rita.desiano@uniparthenope.it

ABSTRACT

The paper seeks to evaluate whether and to what extent the 2007 global financial crisis and its economic aftermaths had any impact on the dynamics of the European countries' national welfare provisions. Relying on available data on 16 Western countries, covering the period 1990-2013, the analysis examines the behaviour of per capita levels of total social expenditure and its main functions: old age-survivors-incapacity related, labour and healthcare. The empirical analysis reveals the presence of a strong conditional convergence process of per capita total public social expenditure and that devolved to functions attracting most of the social resources. The 2007 financial crisis did not change the old age-survivors-incapacity and labour policies' convergence trend while it contributed to increase differences among national indicators of total social provision and health sector with the latter presenting a turnaround with respect to the period before the crisis. Healthcare is confirmed to be a less productive sector and therefore the one that most may be cut in presence of resources availability constraints.

Keywords: convergence, social protection, welfare state indicators, European countries, financial crisis

Study funded by the Unione europea- Next Generation EU, Missione 4 "Istruzione e Ricerca" Componente C2 CUPI53D23002850006
PRIN 2022 "The role of skill- biased technological change in immiserizing growth, spatial polarization and product quality upgrading" – 2022ZKA4PZ

1. Introduction

This study addresses the impact of the 2007 financial crisis on the convergence of welfare policies across European countries. Indeed, up to now the impact of the global financial crises on social protection outcomes has received little attention in the literature, compared to other factors.

The socio-economic changes occurring in developed countries in recent decades been widely investigated (Korpi, Palme, 2003; Montanari, 2001; Tanzi, 2002; Rodrik, 1998; Iversen and Cusack, 2000; Taylor-Gooby, 2004). Most scholars have focused on the effects of the economic slowdowns in the 1970s, the demographic transition, the increase in labour market flexibility, gender equality and higher female labour participation and, finally, on the globalization process leading to a more competitive economic environment. These studies contributed to highlight a new and weaker socio-economic context, with different social risks and priorities that, in the aftermaths of a deep economic and financial crisis, make a far-reaching recalibration of existing welfare policies ever more urgent.

There is a broad consensus in the literature on the role played by the economic crisis in the development of welfare systems. In general, the crisis has been acknowledged to be a crucial breaking point that gives policymakers the strength to divert from their original path and initiate a radical reform process (Vis et al., 2011; Palier, 2010). In "normal" times, indeed, institutional and political forces work against substantial reforms that may question pre-existing rights and privileges and, therefore, would damage in some way their supporters. By contrast, when socio-economic difficulties undermining certainties about future stability that are widespread and perceived as systemic, heavy retrenchment and radical restructuring of welfare programmes appear to be more acceptable and even desirable (Béland and Cox, 2010; Hemmerijck *et al.*, 2009; Kuipers, 2006; Stiller, 2010).

The 2007 global financial crisis, considered as one of the most severe crises of the last century, made most European governments resort to large rescue packages to bail out the financial sector. This had considerable effect on public budgets leading them to worrying increases in deficits. The financial turmoil turned rapidly into an economic crisis and governments' attempts to sustain aggregate demand put additional pressure also on "social insurance" balances (Thalassinos and Politis 2011).

These events led the European Commission to ask Member States to adopt political decision to reduce deficit and debt ratios, to increase employment rates and resume welfare reforms. In particular, to avoid asymmetries in terms of social policies outcomes (Hemerijck and Vandenbroucke, 2012), caused by the lack of a common vision on the goals of a public social support, European institutions expressed the need to achieve a European Social Model (ESM). The main aim of the ESM would be to 'bind Europe together', ensuring that citizens of each Member State feel themselves equally protected, regardless of their country of residence. The revision of welfare objectives and tools would have meant starting with different economic and social contexts and ending up with a situation in which citizens, wherever they live, perceive the same degree of assistance from the state. To this end, European policymakers had at their disposal the tools provided by the Open Method of Coordination. It was a primary mechanism of coordination already employed as part of the European Employment Strategy (EES), in the field of labour market policies, and as a primary instrument in the 2000 Lisbon Strategy, in the fields of poverty, social exclusion, pensions and health care.

That said, this paper seeks to evaluate whether and to what extent the 2007 global financial crisis and its economic aftermaths contributed to achieving some convergence across European countries' national welfare provisions. Relying on available data on 16 Western countries, covering the period 1990-2013, the analysis examines the behaviour of per capita levels of total social expenditure and its main functions. We consider pension, labour and healthcare policy provisions because, taken together, they attract most of the financial resources devoted to social policy and show a larger responsiveness in the turning points of cyclical phases. The analysis seeks to test whether, after controlling for the specific contribution of internal and external conditional factors, the financial crisis played a role in the convergence process of per capita social spending indicators.

The results reveal the presence of a strong conditional convergence process across European countries which concerned all types of social expenditure. The outbreak of the 2007 financial crisis did not affect OSI and labour policies' convergence trend while it contributed to differentiate the national indicators of total social provision and of health sector. This outcome may be explained by the minor role that health expenditure plays in supporting demand and economic activity levels comparing to other social sectors. For this, it represents the sector that can be most affected by cuts in the case of a contraction of resources availability. Health spending began to diverge after the crisis as national economies were hit differently and governments were asked to reorganize the offer of health services according to their own differentiated economic difficulties.

The rest of the study is structured as follows: section two presents the theoretical framework concerning welfare policies in times of crisis and the possible effects of the financial crisis on welfare programme convergence; section three describes the methodology, the dataset and the findings from the descriptive analysis; section four presents the econometric results and section five concludes.

2. Theoretical framework

Changes in welfare provision in Europe have been widely investigated in the comparative welfare state literature. Above all, European economic integration raised interest in understanding how social protection benefits have changed for each EU Member State. However, up to now, analyses investigating the impact of the 2007 financial crisis on European social spending dynamics are scarce or mainly descriptive.

2.1 Welfare policies in times of crisis

What happens to welfare policies in general after a crisis? The literature states that the evolution of welfare programmes most depends on the possibility of a breaking event like an economic crisis turning into a recession (Stark *et al.*, 2014; Vis *et al.*, 2011; Palier, 2010). Usually, following the classical Keynesian intervention strategy, immediately after a crisis, governments are induced to increase resources allocated to welfare policies in order to support the demand sector (Vis *et al.,* 2011; Hermann, 2014). This choice, however, generates considerable stress on public budgets and, when a crisis spreads to the labour market, it contributes to undermining the sustainability of social budgets. Indeed, while the greater incidence of social risks and needs does increase social expenditure, there is a reduction in contribution revenues due to lower economic growth, rises in unemployment and smaller wage increases, which drastically limits the availability of resources. Finally, in the event of a prolonged recession, governments are usually forced to correct deficits and debt ratios by introducing austerity measures, which consist in public spending cuts and/or increases in taxes and contributions (Bonnet et al., 2010; Busch, 2010).

The global financial crisis started as an acute crisis hitting the banking system and later quickly affected the real economy, causing a substantial slump in business investment, household demand and aggregate production. As a first consequence, the governments of most industrialised countries devoted resources to supporting or bailing out banks in an attempt to avoid the collapse of the financial and industrial sectors and prevent a massive drop in domestic demand. These interventions exerted even more pressure on public budgets. Despite such efforts, the financial crisis caused an economic downturn that turned into recession, with negative social and economic consequences. The expansion of social programmes applied new pressure on the sustainability of the pre-crisis social schemes that were already made vulnerable by aging populations, structural changes in the labour market and an increase in female labour market participation that affected the family's very stability.

The increase in social spending together with the strong reduction in taxation revenues and the contamination of sovereign debt by toxic financial sector assets generated broad agreement on the need to consolidate countries' public budgets by adopting appropriate austerity measures (Busch, 2010; Diamond and Lodge, 2013; Hermann, 2014).

Social spending reform appears to be one of the main aims of austerity programmes. The majority of countries have reviewed their welfare schemes, modifying social benefits differently, according to their own inherent conflict between long-term financial sustainability concerns and the counter-cyclical role of welfare spending (Bonnet *et al.*, 2010).

While substantial differences are in general expected when comparing different countries, adjustments in welfare goals in a given countries usually follow change of political orientation due to the alternation of different political parties or the existence of a two-way relationship between policy and public opinion (Brooks and Manza, 2006; Christian, 2008; Kenworthy, 2009; Blekesaune and Quadagno, 2003; Jaeger, 2009). Vis *et al.* (2011) focusing on a sample of countries with different welfare regimes, found that after the recent financial crisis this two-way relationship has worked against rather than opened opportunities for welfare provision retrenchments. Diamond and Lodge (2013), instead, demonstrated that even in the aftermath of a crisis, it resulted very hard to reshape welfare states because of the large inequality in electoral participation found in the majority of advanced democracies. Indeed, the decline in voter participation among the young and the poorest favoured the influence of older voters in the political process, making pension and welfare payments to the elderly almost untouchable.

As regards the differences in policy decisions due to the political composition of governments (Hicks, 1999, Huber and Stephens, 2001, Castles, 1982), Social Democratic parties are usually more likely to favour an expansion of welfare generosity during economic crises with respect to Liberals (Ahrend *et al.*, 2011). Stark *et al.* (2014), finally, found that the influence of parties in shaping responses to the blow of a crisis depends on pre-existing welfare state configurations. Political ideologies may play a role in less generous welfare states while, in more generous welfare states, with the predominance of automatic stabilizers, there is a cross-party consensus concerning the overall direction of social policy changes.

2.2 The financial crisis in the European context

The 2007 global financial crisis contributed to worsen the already precarious conditions of most European countries' public balances. In particular, countries facing serious insolvency problems were allowed to rely on external funds to avoid state bankruptcy, in line with an agreement signed with the European institutions (the European Commission, European Central Bank, European Parliament and the International Monetary Fund). These countries were forced to follow austerity policies aimed at reducing their deficit spending (Busch, 2010; Bieling, 2011; European Commission, 2012a). However, the pressure from the European Commission (EC) to reduce public expenditure was large even for the healthiest countries. As stated by OECD (2012), the overall scale of consolidation measures adopted after the crisis was undoubtedly the most onerous and extensive in Europe since World War II.

The dictates by supranational institutions helped national governments to overcome the resistance to implement the necessary unpopular reforms in the field of social provision. In addition, the presence of common problems (public debt, the unemployment rate, competition due to wage differentials and banking sector stability) and a widespread neoliberal political ideology were able to play a crucial role in steering European countries along a path leading to the convergence of welfare states.

What may have contributed to the convergence is the greater competition due to economic globalisation. In particular, the intention to gain competitiveness may have led to a reduction of the amount of resources devoted to welfare policies, in favour of more productive sectors (Montanari, 2001; Tanzi, 2002), pushing countries towards a downward convergence (Busemeyer, 2009; Garrett and Mitchell, 2001).

Other factors enhancing convergence in European social policies are the learning processes driven by the Open Method of Coordination (Mosher and Trubek, 2003; Trubek and Trubek, 2005) and the legal harmonisation, above all in the case of EMU member states. This harmonisation is achieved through a positive as well as a negative integration (Scharpf, 1998). The former refers to the incorporation of common rules and directives on different aspects of social policy into the national legal framework (Knill and Lehmkull, 1999). Negative integration, instead, refers to the removal of barriers to competition, facilitating the development of the Common Market (Leibfried and Pierson, 1995). This harmonisation enables institutional stickiness to be overcome, acknowledged by the literature (Pierson, 2001; Starke et al., 2008) as one of the main causes of persistence of divergence among countries' social policies even when they face similar challenges.

Finally, the presence of differences in wages, social costs and tax burdens may increase the risk of social dumping and the need for political institutions to limit it by means of proper rules for of competition. To achieve this result, even a general agreement on minimum corporate tax rates, minimum wages and welfare provision goals at supranational level (Busch, 2010) may have the effect of enhancing the convergence of social policy outcomes.

3. Methodology and data

3.1 Methodology

This study investigates the impact of the 2007 financial crisis on the convergence dynamics of welfare state programmes in Western Europe countries. More precisely, following the hypothesis that steady-state levels of welfare provision are strongly influenced by countries' specific characteristics (Alsasua et al., 2007; Attia and Berenger, 2007; Caminada et al., 2010; Paetzold, 2013; Schmitt and Starke, 2011; Starke et al., 2008), we estimate a conditional convergence model following an appropriate panel data approach. The dataset includes 16 countries, namely Austria, Belgium, Denmark, Finland, France, Ireland, Italy, Germany, Greece, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom, and covers the period 1990-2013.

The empirical model we estimate is the following:

$$\Delta SP_{i,t} = \alpha_i + \beta_1 SP_{i,t-1} + \beta_2 DFC + \beta_3 DFC * SP_{i,t-1} + \beta_4 EMU_{i,t} + \gamma_k \pi_{i,t} + \varepsilon_{i,t} \quad \text{with } k=1,\dots 9 \qquad (1)$$

where $\Delta SP_{i,t}$ indicates the annual growth of the social provision indicator of country i (with i=1,…,16), at time t (with t=1, …, 24). $SP_{i,t-1}$ is the lagged value of the social indicator and the coefficient β_1, depending on its sign, reveals the presence of convergence (if negative) or divergence (if positive) among countries' welfare provisions. Two are our variables of interest. The first is a binary dummy variable D*FC*, equal to 1 from 2007 onwards and zero otherwise, that enables us to determine whether and to what extent the global financial crisis has modified the trend of individual countries' social provision indicators. The second variable is given by the interaction term $DFC*SP_{t-1}$ which captures the impact of the crisis on the convergence process of social expenditures. In particular, if the interaction term coefficient (β_3) is statistically significant, it may either reinforce the convergence process, when negative, or weaken it, when positive. The variable EMU is a binary dummy variable that is equal to 1 if the country i in the year t is an EMU member state and zero otherwise the vector $\boldsymbol{\pi}_{i,t}$ includes a parsimonious set of variables controlling for demographic, economic and institutional factors suggested by the reference literature. Finally, specific time-invariant characteristics and structural differences are captured by country-fixed effects α_i.

The use of panel data might present problems of non-stationarity that could increase the probability of undermining the validity of hypothesis tests on the regression parameters or, on the other hand, ending up with an infinite persistence of shock effects. To control for this potential source of bias, we run different unit root tests, namely Levin et al. (2002), Im et al. (2003), ADF Fisher $\chi 2$ and Fisher-PP tests defined by Maddala and Wu (1999), where the null hypothesis is 'non-stationarity' of data. The results of unit root test indicate whether or not, for the explanatory variables included in our empirical model, data need to be differentiated to make them stationary[1].

3.2 Data and descriptive statistics

Following the relevant empirical literature, comparative analyses of welfare programmes usually select social provision indicators on the base of their ability to represent the national welfare scheme. Since this study assesses the impact of the financial crisis on welfare convergence over a limited interval of time (less than ten years after the start of the crisis), we made our choice by looking at the relative size of the social indicator and its responsiveness to the policy changes introduced to combat the aftermath of the crisis itself. On the base of this criterion, we took the per capita public expenditure on overall social services in cash and in kind and the per capita amount of resources allocated to the main welfare functions. In particular, among all functions, we selected the categories attracting most of the financial resources devoted to social policy. To this end, figure 1 gives an idea of the average distribution of social expenditure throughout the sample among different functions in 1990 and 2013. Old age, Survivors, and Incapacity-related (OSI), Health and Labour policies are the most important categories of social spending and, together, they cover about 90% of the total amount of resources available for each citizen in a given country. Data on social expenditure are taken from the OECD Social Expenditure Database (SOCX).

[1] The results of the unit root tests are available on request.

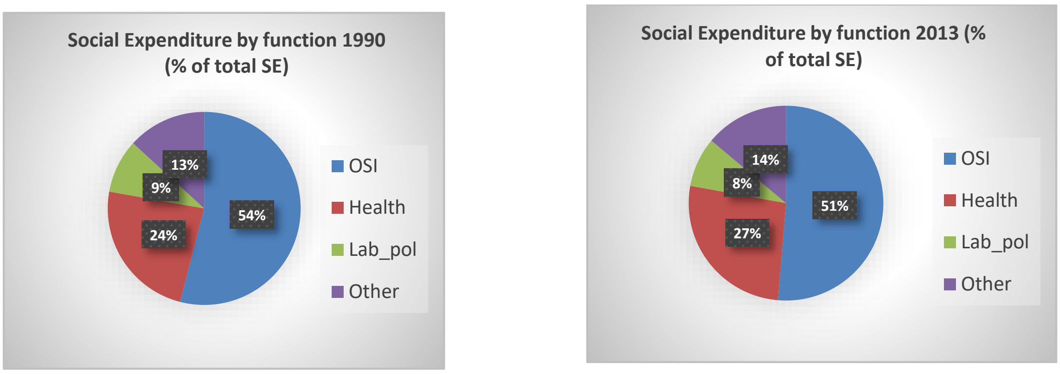


Figure 1. Sample average of per capita social expenditure by function in 1990 and 2013 (our elabouration on OECD-SOCX data)

The literature presents other measures of social provision, such as pension and unemployment replacement rates, particularly helpful in measuring the social policy effort in terms of individual level of social benefits. Nevertheless, these indicators may have some limitations (Otto, 2017; Schmitt and Starke, 2011). They identify the average benefit level calculated on the base of average net wages and are thus strongly influenced by fiscal reforms and wage agreements. Moreover, these indicators do not properly account for other major benefits like housing subsidies and health care. In any case, it would be implausible to expect that well-entrenched programmes like pension or unemployment schemes could be rapidly adapted in response to a crisis. Hence, this type of indicator would not be particularly helpful in our analysis. Similarly, the use of aggregate levels of social expenditure in comparative and convergence studies may present some shortcomings. In particular, it would not allow us to distinguish whether an increase in the indicator would be due to a larger number of beneficiaries (ageing population in the presence of demographic changes or higher unemployment rates during recessions) rather than to changes in policy choices.

To account for these critical issues, we employ social indicators defined in terms of per capita expenditure. Secondly, different demographic, economic and institutional controls are included in the empirical model to eliminate potential biases (Kühner, 2007). Demographic factors are accounted by the old dependency ratio (*old_dep*), i.e. the ratio of the population older than 64 to the working-age population (those aged 15-64), and the crude birth rate (*Crude_birth_rate*). These factors may drive changes in social expenditure for pension, health and education services for children and young people. More precisely, an increase in the old dependency ratio would favour social provision for pension and health services, while a higher birth rate would increase resources allocated to health, family and short-term educational programmes. The pressure on policymaker to increase social spending, especially for family needs like childcare and eldercare, increase also when there are higher levels of female participation in the labour force (*femalpart*) (Ennser-Jedenastik, 2017; Andreotti et al. 2013). A larger entry of women into paid work may cause a weakening of the "male breadwinner" model leading so to a more homogeneous social configuration across European countries that start from a very different gender division of tasks both inside and outside the family (Esping-Andersen, 2009).

Among the economic controls, we include the GDP growth rate (*tx-Gdp*) which, by improving social well-being, favours a reduction in the demand for resources devoted to social functions. Completely different could be the behaviour of healthcare spending. In this case, an increase of wealth should lead to a rise in the demand for health goods and services due to the greater attention of individuals for their own health. An increase of the unemployment rate (*tx_unemp*), instead, requires a higher disbursement in different social spending functions (labour, health, family and housing). The international openness is captured by the trade openness (t*rade_open*), that is the sum of imports and exports in terms of GDP, and an index of restrictions on cross border capital transactions (*kaopen*) ranging between 0 (minimal openness) and 1 (maximal openness). The latter two indicators, in a context of higher globalization, enable to measure the exposure of a country to external risks and to interpret the government responses in the light of the efficiency vs. compensation views in the theoretical debate. The efficiency hypothesis implies a strong reduction of social spending in order to strengthen a more productive use of public resources and gain in terms of competitiveness. The compensation hypothesis (Avelino et al., 2005), instead, states that governments act to protect domestic interests and mitigate the inequality associated with openness through the strengthening of mechanisms for social insurance (social security, pensions, unemployment insurance).

Finally, we include some controls for institutional features of European countries that, above all for EMU member states, appear to be particularly conditioning. The recent changes in fiscal governance at the EU level, for EMU countries in particular, left little space for a completely independent decision-making process so that accounting for this membership is essential (*EMU*). Supranational limits result to be added to those already existing at national level in terms of rationalization of public balances. To this end, we include government debt in the percentage of GDP (*debtgdp*) and we expect that, in the perspective of public spending tightening, and increase in public debt drives to an even larger restriction of social expenditures, in general acknowledged as unproductive. To conclude, we include a control for the left wing party (*govleft1*), to account for the tendency of Social democratic parties to favour an expansion of welfare spending during economic crises with respect to right wing parties

(Ahrend *et al.*, 2011). Left wing parties show strong political preferences for labour policies, probably due to their origin in the 19th century from the labour movement. Adams *et al*. (2009), in particular, explain the restrictions they usually show in ideological flexibility with the strong link they keep with trade unions and social movements (Kitschelt,1994; Piazza, 2001, Bremer, 2018).

At the end of this paragraph, we present the results of a brief descriptive analysis of the social provision highlighting the differences between countries. Altogether, the 16 European countries show increasing annual average values of the four indicators throughout the interval period (Figure 2).

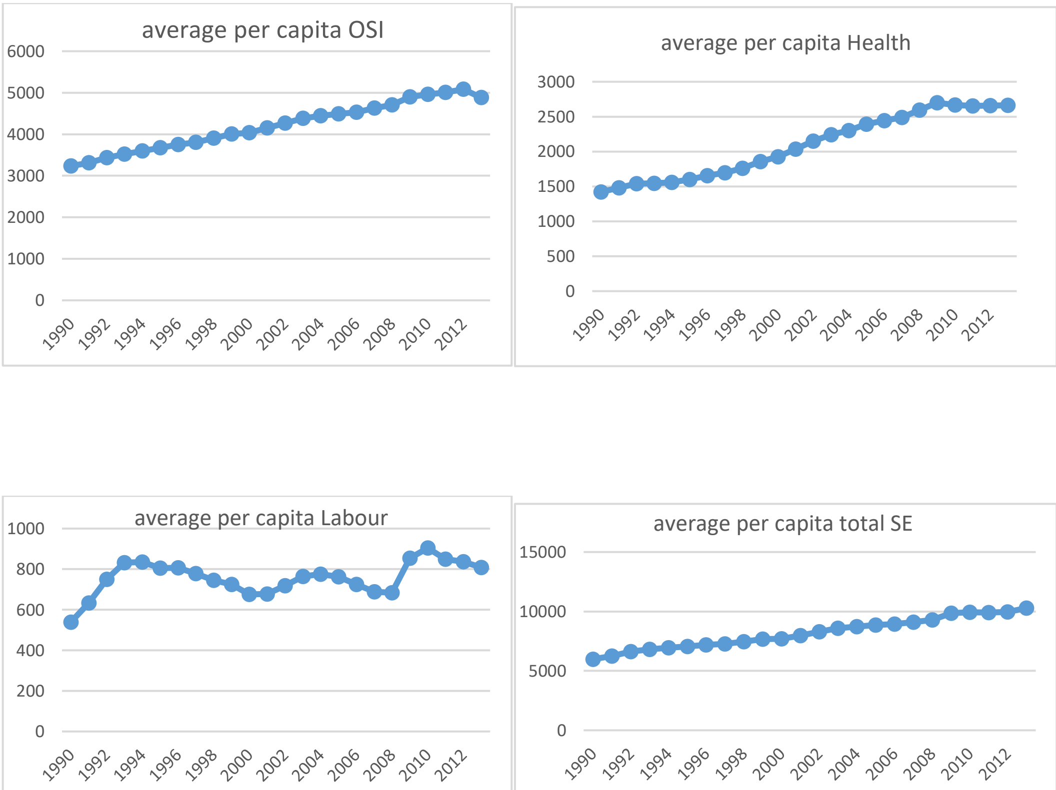


Figure 2. Average annual social indicators for the 16 European countries

The 2007 financial crisisdoes not seem to have had a noticeable effect on the upward trend of social provisions, with the sole exception represented by the per capita amount of resources allocated to labour policies. This evidence may be explained by the introduction of measures aimed at managing new socio-economic risks, driven by the dramatic pressures in particular on the labour market.

Data in Table 1 enable to focus on changes at country level. The general increase in per capita social benefits from 1990 to 2013 shows greater generosity for Portugal, Ireland, Greece and the United Kingdom and less for the Netherlands, Germany, Sweden and Norway. The increase in the share of total social spending allocated to OSI is higher in Portugal, Ireland Greece and Finland and lower in the Netherlands, Germany, Norway and Sweden. The distribution of social resources devoted to health benefits increases most in Ireland, the United Kingdom, Switzerland and Greece, and less in Italy, Finland, Sweden and Belgium. Finally, labour policy absorption of social resources rises most in Switzerland, Portugal, Denmark and Italy while it even decreases in some countries like the UK, Norway and the Netherlands. Therefore, the suggestion is that countries have increased public social spending, to limit the consequences of the financial crisis on individuals and families, by means of automatic stabilizers and discretionary measures aimed at reinforcing social benefits.

When looking at the years immediately after the crisis, data confirm this hypothesis. Analysis of per capita social indicators annual growth rates reveals a temporary increase in total social spending and the other main functions with the only exception of health spending (Table 2: a, b, c and d). This may be explained by the role played by social policies in supporting and preventing risks that are particularly burdensome in times of crisis. The increase in expenditure on labour policies was particularly large (5.41% on average) especially in Nordic and Continental countries, where higher unemployment protection most likely was implemented through the automatic stabilizers.

Table1. Changes in per capita total and sectoral social expenditures in the period 1990-2013

Country	SE	OSI	Health	Labour
Austria	68.14	63.81	84.06	102.47
Belgium	60.81	55.83	69.84	35.93
Denmark	70.12	59.51	93.79	217.52
Finland	72.46	83.09	40.30	110.96
France	63.76	60.37	79.73	34.26
Germany	40.91	29.51	59.33	22.32
Greece*	111.26	104.55	126.49	175.07
Ireland	159.78	135.16	187.94	136.81
Italy	51.21	48.73	30.24	188.14
Netherlands	33.90	4.60	118.69	-3.04
Norway	48.06	40.62	109.82	-35.50
Portugal	165.65	183.52	115.32	253.63
Spain	79.67	88.66	74.92	27.59
Sweden	41.32	43.01	59.50	7.94
Switzerland	88.78	51.61	134.57	423.16
United Kingdom	103.07	79.03	138.86	-35.96
16 European countries	**78.68**	**70.73**	**95.21**	**103.83**

* Data for Greece are referred to years 1990-2012.

Average resources devoted to pension provision increased after the crisis to mitigate the potential social risks caused by the recession. Rises in pension spending may occur in spite of the negative economic situation undermining the financial sustainability of pension schemes. European governments have tried to tackle the sustainability problem through reforms that might take effect, however, only in the medium or long term (well beyond the considered period).

Health spending shrank between 2009 and 2010. This is the only function in which the cut in resources, introduced through annual healthcare plans, showed any effect in the short term.

Table 2a. Annual growth in per capita total social spending

Country	*2009/2010*	*2010/2011*	*2011/2012*	*2012/2013*
Austria	1.88	-0.39	1.77	1.04
Belgium	0.71	2.50	0.42	0.59
Denmark	3.38	0.45	-0.28	-0.26
Finland	4.34	1.00	2.71	2.57
France	1.87	0.93	1.56	1.71
Germany	1.42	-1.37	-0.19	0.82
Greece	-5.17	-1.00	0.71	
Ireland	0.95	-4.24	-0.17	-2.40
Italy	1.00	-0.87	-0.59	-0.42
Netherlands	3.37	0.66	1.01	0.83
Norway	-2.29	-2.64	1.06	1.77
Portugal	1.44	-2.15	-3.00	3.34
Spain	1.42	0.39	-3.42	-0.63
Sweden	-0.18	0.01	2.52	2.95
Switzerland	0.96	0.45	2.57	2.84
United Kingdom	-0.01	-0.59	1.04	-1.15
Average	**0.94**	**-0.43**	**0.48**	**0.91**

Table 2b. Annual growth in per capita OSI spending

Country	*2009/2010*	*2010/2011*	*2011/2012*	*2012/2013*
Austria	2.52	0.99	2.11	1.26
Belgium	1.14	3.11	1.04	2.90
Denmark	1.51	1.12	0.58	0.21
Finland	4.48	1.73	3.13	2.68
France	1.97	2.64	1.86	1.70
Germany	1.46	-0.06	0.57	-0.26
Greece	-1.19	-0.60	9.31	
Ireland	4.07	-1.16	0.90	-1.33
Italy	2.22	-0.23	-0.20	0.19

Netherlands	4.42	1.77	0.92	1.82
Norway	-3.95	-2.01	2.30	2.69
Portugal	2.24	3.05	-2.52	6.50
Spain	4.33	2.06	-0.20	2.97
Sweden	-2.03	-1.14	2.71	3.20
Switzerland	0.06	2.23	0.63	0.90
United Kingdom	-0.08	0.91	2.18	-0.20
Average	**1.45**	**0.90**	**1.58**	**1.68**

Table2c. Annual growth in per capita Labour spending

Country	*2009/2010*	*2010/2011*	*2011/2012*	*2012/2013*
Austria	-1.37	-7.94	0.17	6.54
Belgium	1.30	1.86	-2.40	-6.95
Denmark	26.70	-0.45	-5.57	-2.07
Finland	9.27	-8.91	1.49	7.24
France	8.85	-9.42	1.32	1.86
Germany	-5.58	-16.66	-10.05	-0.73
Greece	9.07	13.31	-7.94	
Ireland	23.18	-6.80	-2.57	-6.49
Italy	-2.75	1.09	4.81	0.80
Netherlands	6.04	-5.78	3.90	-3.34
Norway	8.88	-12.88	-8.67	-3.44
Portugal	1.45	-8.51	10.64	3.03
Spain	-2.66	2.31	-9.90	-6.37
Sweden	9.46	-3.95	8.65	5.71
Switzerland	4.55	-20.78	1.54	8.82
United Kingdom	-9.89	-17.54	-0.98	-12.65
Average	**5.41**	**-6.32**	**-0.97**	**-0.54**

Table 2d. Annual growth in per capita Health spending

Country	2009/2010	2010/2011	2011/2012	2012/2013
Austria	0.79	0.42	2.48	-0.55
Belgium	-0.27	2.27	1.29	0.50
Denmark	-0.46	-0.55	0.97	-1.29
Finland	2.52	3.17	2.23	1.52
France	0.55	1.41	0.86	1.51
Germany	2.78	0.57	0.68	2.29
Greece	-13.46	-2.17	-12.49	-5.73
Ireland	-12.45	-5.86	1.69	0.90
Italy	1.19	-2.96	-3.16	-2.29
Netherlands	2.39	-0.10	3.76	-0.75
Norway	-1.94	-2.71	1.87	0.90
Portugal	0.58	-7.84	-8.06	-1.94
Spain	-0.83	-1.84	-5.03	-3.92

Sweden	-0.40	2.47	0.65	1.69
Switzerland	0.83	2.79	5.03	4.26
United Kingdom	-1.65	-1.58	-0.49	0.56
Average	**-1.24**	**-0.78**	**-0.48**	**-0.15**

The data in Table 2 clearly indicate that both health and labour expenditures on average diminished after the global financial and economic crisis (from 2010 onwards). Greece, Italy and Portugal show a large reduction in per capita health spending in 2013. The drop was extremely dramatic for Greece immediately after the crisis, amounting to -13.5% in 2010 and -12.5% in 2013. As regards Portugal and Italy, they showed tightening health spending for three consecutive years. Austria and the Netherlands posted real-term drops in health spending for the first time in 2013.

4. Results

Empirical results showed in Table 3 reveal the presence of a strong conditional convergence for all the social indicators. Indeed, the corresponding coefficients, the *βs,* are always negative and highly statistically significant (at the 1% level of significance) which testifies the presence of a convergence process.

The binary dummy variable *DFC* enables to verify whether and to what extent the global financial crisis has modified the trend of individual countries' social provision. Estimation results show that the crisis played an important and statistically significant role only on per capita healthcare spending. A possible explanation is the fact that, during and after the crisis, countries faced an objective difficulty in keeping constant the amount of resources devolved to social issues and, among them, healthcare may be considered as the one that supports the aggregate demand less than others. Therefore, it results the most suitable for cuts in the case of resource scarcity making so more difficult to plan investments and provide goods and services in the health sector. This outcome is in line with the previous literature that acknowledged the global financial crisis as also a health system shock (Mladovsky *et al.*, 2012). No significant effects can be found of the crisis on the annual growth for the remaining social indicators.

The interaction term captures the impact of the global financial crisis on the social provision convergence process. The crisis appears to have undermined the convergence processes of total social spending and health spending while it had no effects on OSI and labourlabour. This could be due to the presence of major constraints imposed by the European Union’s social policy, aimed at promoting employment, adequate social protection and in doing so mitigating problems of exclusion. On one hand, in fact, the European Employment Strategy in 1997 and the Lisbon Strategy in 2000 were aimed at overcoming the critical issues of labour market. On the other hand, the EC with the Green Paper on pensions (EC, 2010) pointed out the importance on a common concern about pension systems sustainability already endangered by the ongoing European demographic trends. In this regard, the main suggestion was to start building up safe complementary retirement savings to prevent public balances risky conditions due, in particular, to the retirement of the so called "baby-boomers" and the shrink of working age population. The European Parliament, the European Economic and Social Committee and the Committee of the Regions took part to the debate asking for a comprehensive and coordinated strategy developed at the EU level. The EC response came through the publication of the White Paper '*An Agenda for Adequate, Safe and Sustainable Pensions*' (EC, 2012b) aimed at suggesting forward policy orientation to support national policymakers in their effort to address reform needs.

In truth, estimation results reveal that European Union membership, controlled by the dummy variable *EMU*, plays a decisive role for the overall social spending sectors. Coefficients for all the models are highly significant and positive except for labour policy spending. This positive relationship is probably due to a stronger coordination of social policy by national institutions of EMU members. The EU social policy framework highlights three different concerns. From a social point of view, all countries should accept the responsibility for the social needs of their citizens and agree with the idea of a common European Social Model (ESM). Second, from an economic perspective, countries have to promote competition and avoid distortions due to differences in social conditions caused by discrimination concerning education and labour mobility. Finally, the political concern refers to the presence of an EU active social policy as a *conditio sine qua non* to obtain citizens' consensus on the political and economic integration (Bonasia and De Siano, 2017).

Another major result of this study concerns the role played by demographic, economic and institutional factors. When statistically significant, these control variables usually have the expected signs. This is the case of the GDP annual growth rate, which shows a negative effect on social indicators, meaning that an improvement in social well-being conditions may favour a contraction of resources devoted to per capita social benefits. Likewise, an increase in the unemployment rate has a positive and highly significant impact on the per capita disbursements arranged by labour policies (2.281). A rise in the crude birth rate increases significantly, as expected, total social expenditure and labour spending. Particularly interesting is the evidence on left wing parties influence, which is positive and significant only in the model for the labour policy expenditure. This result confirms the tendency of Social democratic parties to protect the unemployed.

What requires a deepening is the presence of unexpected effects which conflict with previous findings. First, the negative impact of unemployment on health spending (-0.863) could be attributed primarily to budgetary problems: the increase in unemployment causes a reduction in health insurance contributions and hence cuts and/or reorganizations of resources available for heath policies. Confirming this hypothesis, a report of OECD Health Statistics (OECD, 2015) underlines that *"…the need to reduce public deficits in some countries has resulted in a shift to private sources of financing via changes to entitlement, amendments to the benefits package and the introduction/extension of user charges*". Secondly, the negative impact of a larger female labourlabour force participation on total social spending and on pension and labour provisions policy spending. The higher is women participation in the labour force the less the strength by government in sustaining both individuals and families and, hence the amount of resources employed for poverty reduction, income maintenance, employment enforcement and gender equality (Gehringer *et al*., 2014). Finally, the absence of statistically significant impact of old age dependency ratio on OSI may be explained, in a context of legal harmonisation, as the consequence of EU general warnings to national governments to maintain an adequate standard of living and to promote private supplementary pension schemes (EC, 2012b).

Table 3. Panel data estimation results

	(1)	(2)	(3)	(4)
VARIABLES	Tot_Soc_exp	OSI	Labour	Health
L.SP	-0.00174***	-0.00296***	-0.0234***	-0.00433***
	(0.000242)	(0.000451)	(0.00369)	(0.00113)
DFC	-2.663	0.795	-0.0446	-8.681***
	(1.802)	(1.387)	(3.305)	(3.054)
DFC*SP	0.000400**	9.33e-05	0.000626	0.00335***
	(0.000179)	(0.000270)	(0.00365)	(0.00118)
DIFF_debtgdp	0.000522	0.00445	0.00444	-0.0110*
	(0.00341)	(0.00325)	(0.0138)	(0.00579)
DIFF_tx_unemp	0.142	0.0573	2.281***	-0.863***
	(0.167)	(0.158)	(0.673)	(0.285)
tx_Gdp	-0.257***	-0.180**	-0.943**	-0.307*
	(0.0925)	(0.0890)	(0.375)	(0.158)
Crude_birth_rate	0.527**	0.0857	4.000***	-0.344
	(0.217)	(0.212)	(0.880)	(0.370)
trade_open	0.000660	0.00563	-0.191**	-0.00525
	(0.0194)	(0.0177)	(0.0782)	(0.0325)
old_dep	-0.0701	-0.0585	0.566	-0.415**

	(0.112)	(0.108)	(0.465)	(0.196)
govleft1	0.00267	0.000872	0.0398**	-0.00471
	(0.00417)	(0.00400)	(0.0171)	(0.00708)
DIFF_femalpart	-0.441**	-0.429**	-3.053***	0.683**
	(0.182)	(0.174)	(0.739)	(0.311)
kaopen	1.154	2.704*	-8.454	-0.235
	(1.445)	(1.414)	(5.328)	(2.428)
EMU	2.659***	1.896***	3.337	4.062***
	(0.560)	(0.530)	(2.209)	(0.983)
Constant	10.21**	11.31**	-14.91	25.13***
	(4.755)	(4.561)	(19.03)	(8.212)
Observations	362	362	362	363
R-squared	0.283	0.184	0.415	0.204
Number of Country	16	16	16	16

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

5. Conclusions

This study analysed the impact of the 2007 global financial crisis on European countries' national welfare provisions to verify whether it may have had any influence on their convergence process.

European welfare states have a high degree of institutional diversity due to political, historical and economic antecedents, and belong to different welfare regimes, in terms of structure, types of risks and needs covered and provision of resources. At the same time, they share the same difficulties in terms of unemployment and aging population. This common stimulus may have prompted policymakers to identify similar strategies in response to similar problems (Adelantado and Calderón Cuevas, 2006). In light of these considerations, the paper sought to answer the following questions. By controlling for individual country demographic, economic and institutional characteristics, do European countries converge in terms of welfare policy provisions? Did the financial crisis contribute to reinforce or to weaken this process with regard to total social expenditure and its main functions.

The empirical analysis revealed the presence of a strong conditional convergence process for total public social expenditure and all the functions attracting most of the social resources, namely OSI, health and labour policies. However, our main outcome concerns the evidence that the onset of the financial crisis did not change neither OSI and labour policies' convergence trend while it contributed to differentiate the national indicators for the total social provision and the health expenditure, both displaying a turnaround with respect to the period before the crisis. Changes in spending programmes introduced in the 1990s in the field of pensions and labour policies managed to stem the negative aftermaths of the economic crisis by lending substantial support to domestic aggregate demand. By contrast, different approaches to healthcare spending were found among European countries. The economic literature, in this regard, argues that health spending is the one that can be most affected by cuts in the perspective of reducing unproductive expenses in presence of a contraction of resources availability. The increase in health spending heterogeneity strategies, therefore, may be explained by differences in economic pressures and in national governments' reactions in terms of budget resources reallocations and/or sensitivities to public assistance in times of crisis.

The absence of any significant impact of the financial crisis on the convergence process of labour and pension function indicators, instead, could be explained by the presence of a severe constraint imposed by the European Union's social policy supporting employment and an adequate social protection level aimed at combating social exclusion. Indeed, in 2000 both the Lisbon Strategy and the EU Social Inclusion Strategy agreed to achieve the goal of poverty eradication following the Open Method of Coordination. To reinforce these intentions, the Lisbon treaty of 2009 and the more recent Europe 2020 strategy included a "social clause" whereby the above social issues must be considered when defining and implementing whichever policy.

References

Adams, J., Haupt, A. B., & Stoll, H. (2009). What Moves Parties? The Role of Public Opinion and Global Economic Conditions in Western Europe. *Comparative Political Studies*, 42(5), 611–639.

Adelantado, J., & Calderón, Cuevas, E. (2006). Globalization and the Welfare State: The Same Strategies for Similar Problems? *Journal of European Social Policy*, 16(4), 374-386.

Ahrend, R., Arnold, J., & Moeser, C. (2011). The Sharing of Macroeconomic Risk: Who Loses (and Gains) from Macroeconomic Shocks. *OECD Economics Department Working Papers* 877. Paris: OECD Publishing.

Alsasua, J., Bilbao-Ubillos, J. & Olaskoaga, J. (2007). The EU Integration Process and the Convergence of Social Protection Benefits at National Level. *International Journal of Social Welfare*, 16(4), 297–306.

Andreotti, A., Mingione, E., & Pratschke, J. (2013). Female employment and the economic crisis: Social change in Northern and Southern Italy. *European Societies*, 15(4), 617-635.

Attia, N. & Berenger, V. (2007). Social protection Convergence in the European Union: impact of Maastricht Treaty. *Panoeconomicus*, 4(4), 469-87.

Avelino, G., Brown, D., & Hunter, W. (2005). The Effects of Capital Mobility, Trade Openness, and Democracy on Social Spending in Latin America, 1980-1999. *American Journal of Political Science* 49(3), 625-41.

Béland, D. & Cox, R. H. (eds) (2010*). Ideas and Politics in Social Science Research*. Oxford: Oxford University Press.

Bieling, H.J. (2011). Vom Krisenmanagement zur neuen Konsolidierungsagenda der EU. *Prokla*, 41(2), 173–94.

Blekesaune, M. & Quadagno, J. (2003). Public Attitudes toward Welfare State Policies: A Comparative Analysis of 24 Nations. *European Sociological Review*, 19(5), 415-427.

Bonasia, M. & De Siano, R. (2017). Convergence of welfare provision in Europe: An assessment of the Monetary Integration effects. *CRISEI Discussion Paper* n.07, ISSN: 2280-9767.

Bonnet, F., Ehmke, E. & Hagemejer, K. (2010). Social security in times of crisis. *International Social Security Review*, 63, 47–70.

Bremer, B. (2018). The missing left? Economic crisis and the programmatic response of social democratic parties in Europe. *Party Politics*, vol. 24(1), 23–38.

Brooks, C. & Manza, J. (2006). Why do Welfare States Persist? *The Journal of Politics*, 68(4), 816-827.

Busch, K. (2010). World Economic Crisis and the Welfare State, *International Policy Analysis*. Friedrich Ebert Stiftung.

Busemeyer, M. (2009). From Myth to Reality: Globalisation and Public Spending in OECD Countries Revisited. *European Journal of Political Research*, 48 (4), 455–82.

Caminada, K., Goudswaard, K. & Van Vliet, O. (2010). Patterns of Welfare State Indicators in the EU: Is there Convergence? *Journal of Common Market Studies*, 48(3), 529–556.

Castles, F.G. (ed) (1982). *The Impact of Parties: Politics and Policies in Democratic Capitalist States*. London: Sage.

Christian, J.L. (2008). When Does Public Opinion Matter? *Journal of Sociology and Social Welfare*, 35(1), 133-156.

Diamond, P. & Lodge, G. (2013). Welfare States after the Crisis: Changing Public Attitudes. *Policy Network Paper*, www.policy-network.net.

Enner-Jedenastik, L. (2017). How women's political representation affects spending on family benefits. *Journal of Social Policy*, 46(3), 563-581.

Esping-Andersen, G. (2009). *Incomplete revolution: Adapting welfare states to women's new roles*. Polity.

European Commission, (2010). *Green paper: Towards adequate, sustainable and safe European pension systems.* SEC, 830. Brussels.

European Commission, (2012a). *Employment and social developments in Europe*. Brussels, European Commission.

European Commission, (2012b). *An Agenda for Adequate, Safe and Sustainable Pensions: White Paper*. Publications Office of the European Union.

Garrett, G. & Mitchell, D. (2001). Globalization, government spending and taxation in the OECD. *European Journal of Political Research*, 39(2), 145–177.

Gehringer, A., Klasen, S., & Witkowska, D. (2014). Labour Force Participation and Family Policies in Europe: an Empirical Study. Quantitative Methods in Economics, 15(1), 37-49.

Hemerijck, A., Knapen, B., & van Doorne, E. (2009). *Aftershocks. Economic Crisis and Institutional Choice*. Amsterdam: Amsterdam University Press.

Hemerijck, A.C. & Vandenbroucke, F. (2012). Social investment and the Euro Crisis: The Necessity of a Unifying Social Policy Concept. *Intereconomics*, 1-7, ZBW – Leibniz Information Centre for Economics.

Hermann, C. (2014). Structural adjustment and neoliberal convergence in labour markets and welfare: The impact of the crisis and austerity measures on European economic and social models. *Competition & Change*, 18(2), 111–130.

Hicks, A. (1999). *Social Democracy and Welfare Capitalism. A Century of Income Security Politics* New York: Cornell University Press.

Huber, E. & Stephens, J.D. (2001). *Development and Crisis of the Welfare State: Parties and Policies in Global Markets*. Chicago: University of Chicago Press.

Im, K.S., Pesaran, M.H & Shin, Y. (2003). Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics,* 115, 53–74.

Iversen, T. & Cusack, T. (2000). The causes of welfare state expansion: deindustrialization or globalization? *World Politics*, 52, 313–49.

Jaeger, M.M. (2009). United But Divided: Welfare Regimes and the Level and Variance in Public Support for Redistribution. *European Sociological Review*, 25(6), 723-737.

Kenworthy, L. (2009). The effect of public opinion on social policy generosity. *Socio-Economic Review,* 7(4), 727-74.

Kitschelt, H. (1994). *The transformation of European Social Democracy*. New York: Cambridge University Press.

Knill, C. & Lehmkuhl, D. (1999). How Europe Matters: Different Mechanisms of Europeanization. *European Integration Online Papers*, 3(7), http://eiop.or.at/eiop/texte/1999-007.htm

Korpi, W. & Palme. J. (2003). New Politics and Class Politics in the Context of Austerity and Globalization: Welfare State Regress in 18 Countries. *American Political Science Review,* 97, 425–46.

Kühner, S. (2007). Country-level comparisons of welfare state change measures: Another facet of the dependent variable problem within the comparative analysis of the welfare state. Journal of European Social Policy, 17(1), 5-18.

Kuipers, S.L. (2006). *The Crisis Imperative: Crisis Rhetoric and Welfare State Reform in Belgium and the Netherlands in the Early 1990s*. Amsterdam: Amsterdam University Press.

Leibfried, S. & Pierson, P. (1995). *European Social Policy. Between Fragmentation and Integration*. Washington DC: Brookings Institution.

Levin A, Lin C.F. & Chu, C. (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties. *Journal of Econometrics*, 108, 1–24.

Maddala, G.S. & Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and A New Simple Test. *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

Mladovsky, P., Srivastava, D., Cylus, J., Karanikolos, M., Evetovits, T., Thomson, S. & McKee, M. (2012). *Health policy responses to the financial crisis in Europe*. Copenhagen:World Health Organization.

Montanari, I. (2001). Modernization, globalization and the welfare state: a comparative analysis of old and new convergence of social insurance since 1930. *British Journal of Sociology*, 52(3), 469-94.

Mosher, J. & Trubek, D. (2003). Alternative Approaches to Governance in the EU: EU Social Policy and the European Employment Strategy. *Journal of Common Market Studies*, 41(1), 63–88.

OECD, (2012). *Restoring Public Finance*. Paris: OECD.

OECD, (2015). *FOCUS on Health Spending*. OECD Health Statistics.

Otto, A. (2017). A Benefit Recipiency Approach to Analysing Differences and Similarities in European Welfare Provision. *Social Indicators Research*, 137(2), 765-788. https://doi.org/10.1007/s11205-017-1621-5.

Paetzold, J. (2013). The Convergence of Welfare State Indicators in Europe: Evidence from Panel Data. *European Journal of Social Security*, 15(1), 28–55.

Palier, B. (2010). The long conservative corporatist road to welfare reforms. In B. Palier (ed.), A Long Goodbye to Bismarck? *The Politics of Welfare Reform in Continental Europe*, 333–387. Amsterdam: Amsterdam University Press.

Piazza, J. (2001). De-linking labour: labour unions and social democratic parties under globalization. *Party Politics,* 7(4), 413–435.

Pierson, P. (ed) (2001). *The New Politics of the Welfare State*. Oxford: Oxford University Press.

Rodrik, D. (1998). Why do more open economies have bigger governments?. *Journal of Political Economy*, 106(5), 997-1032.

Scharpf, F.W. (1998). Negative and Positive Integration in the Political Economy of European. Welfare States. In Rhodes, M. & Meny, Y. (eds), *The Future of European Welfare*. A New Social Contract?. London: Macmillan Press Ltd., 157-77.

Schmitt, C. & Starke, P. (2011). Explaining Convergence of OECD Welfare States: A Conditional Approach. *Journal of European Social Policy*, 21(2), 120–35.

Starke, P., Kaasch, A., & van Hooren, F. (2014). Political Parties and Social Policy Responses to Global Economic Crises: Constrained Partisanship in Mature Welfare States. *Journal of Social Policy*, 43(2), 225-246.

Starke, P., Obinger, H., & Castles, F.G. (2008). Convergence towards Where: In What Ways, If Any, are Welfare States becoming More Similar? *Journal of European Public Policy*, 15(7), 975–1000.

Stiller, S. (2010). *Ideational Leadership in German Welfare State Reform: How Politicians and Policy Ideas Transform Resilient Institutions*. Amsterdam: Amsterdam University Press.

Tanzi, V. (2002). Globalization and the Future of Social Protection. *Scottish Journal of Political Economy*, 49, 116–27.

Taylor-Gooby, P. (ed.) (2004). New Risks, New Welfare: The Transformation of the European Welfare State. Oxford: Oxford University Press.

Thalassinos, I.E. & Politis, E.D. (2011) International Stock Markets: A Co-integration Analysis. European Research Studies Journal XIV(4), 113-130.

Trubek, D.M. & Trubek, L.G. (2005). Hard and Soft Law in the Construction of Social Europe: the Role of the Open Method of Co-ordination. European Law Journal 11(3), 343–364.

Vis, B., van Kersbergen, K., & Hylands, T. (2011). To What Extent Did the Financial Crisis Intensify the Pressure to Reform the Welfare State? Social Policy & Administration 45(4), 338–353.

Factors influencing the speed of public spending: institutional quality, administrative capacity and decentralization

Alfredo Del Monte[1], Alessandro De Iudicibus[2], Sara Moccia[1,†] and Luca Pennacchio[3]

Abstract

The fast and effective use of public resources is crucial to economic development. This study examines how institutional quality and administrative capacity influence the speed of public spending across different levels of government. Drawing on a comprehensive dataset of public projects co-financed by European funds from 2010 to 2019, the analysis explores both internal and external administrative factors affecting spending speed. The findings underscore the critical importance of administrative capacity and institutional quality, with administrative capacity serving as a mediator for the effects of institutional quality. The speed of public spending is higher for regional and provincial governments than for municipal governments. Furthermore, the estimates suggest that institutional quality has a greater influence at the lower levels of government, whereas administrative capacity plays a more significant role at the higher levels of government.

Keywords: Quality of institutions; Administrative capacity; Decentralization; Public spending; OpenCoesione

JEL Classification: D73, H11, H70, R50

[1] University Federico II of Naples, Department of Economics and Statistics. [†]Corresponding Author: sara.moccia@unina.it
[2] University of Campania Luigi Vanvitelli, Department of Economics
[3] Parthenope University, Department of Business and Economics.

1. Introduction

Historical evidence regarding public projects and investment dynamics highlights Italy's tendency toward a slow speed of public spending. This slowness in expenditure often characterizes the country's approach to public initiatives and investments. Therefore, the duration of public works is particularly high in Italy compared to other European countries (Agenzia per la Coesione Territoriale 2018). Serious concerns have surfaced regarding Italy's ability to plan and utilize European Union (EU) funds effectively. The persistent sluggishness in spending public funds has long been recognized as a weakness in Italian economic policy (Del Monte et al. 2022b).

The effect of the EU Cohesion Policy is greatly affected by the territorial context and public administration capacity of local governments in which such policy is implemented. It is therefore important to understand which factors affect such aspects. Charron (2014) studied the quality of local administrations in more than 200 European regions, and found that the top Italian region was in 118th place. Southern Italian regions were at the lowest levels of this ranking. Putnam et al. (1993) documented disparities in the quality of local administrations between Northern and Southern Italian regions. Del Monte et al. (2022a) examined the impact of government decentralization on the duration of public projects in Italy. They discovered regional disparities in project duration, noting that in the South and certain Central regions, projects took longer at the lowest government level of government but were faster at more centralized levels. This implies that decentralization in these areas decreased the pace of public spending and, consequently, spending efficiency. In contrast, in the North and most Central regions, the findings were inconclusive, with regional differences primarily influenced by the quality of local institutions.

Regarding the impact of administrative capacity on public funding, Bachtler et al. (2023) emphasized that both internal and external administrative factors affect the rate of absorption, regularity and effectiveness of EU Cohesion Policy.[4] Administrative capacity is influenced by policy measures and norms governing public spending. In Italy, there is a debate regarding whether current norms increase the complexity of procedures and thus decrease the speed of public spending. It has been suggested to reduce the number and type of controls. However, a reduction in controls on public spending could compromise the quality of institutions and GDP.

This paper aims to examine whether the quality of institutions and administrative capacity affect the speed of public spending across various levels of government. We consider a large sample of public

[4] More in detail the authors said that the internal components of administrative capacity are variously defined and analysed as: human, organisational structures, physical resources, systems and tools, leadership, openness to external knowledge. Differently to a range of other factors that have also been identified as external to the administrative system for Cohesion Policy but influencing the ability of the system to achieve its objectives, especially: the quality of public administration, legal stability/regulatory quality, centralisation/decentralisation of governance, political influence

projects co-financed by the European Union funds during the period 2010-2019 and managed by various levels of Italian government (municipalities, provinces, and regions) to examine if spending speed depends on both internal and external administrative factors. The duration of these projects is used as an inverse measure of spending speed and therefore efficiency in the use of public resources (Del Monte et al. 2022a). The speed of expenditure plays a fundamental role in assessing the ability of governments to utilize public resources effectively. The swiftness with which funds are assigned and used serves as a necessary gauge of project execution, especially in nations such as Italy, where efficient utilization of public resources is a challenge. The efficient absorption of EU resources, for example, is a prerequisite for the effectiveness of EU regional policy (Bachtler et al. 2018). The empirical results highlight the fundamental role of institutional quality and administrative capacity in shaping spending speed, with administrative capacity mediating the impact of institutional quality. The impact of institutional quality is stronger for the lower levels of government while the impact of administrative capacity is stronger for the higher levels of governments.

This article contributes to the literature in several ways. It adds to the debate on the effectiveness of EU funds, showing that a low quality of local governments managing public projects can have a detrimental effect on the speed of public spending. Previous studies have largely emphasized a positive correlation between the quality of government and the absorption of cohesion funds (Baun &Marek 2017; Milio 2007; Terracciano & Graziano 2016) but only a few studies have investigated how the quality of institutions affects the speed of spending (Del Monte et al. 2022a; Del Monte et al. 2022b). Our paper also contributes to the ongoing debate regarding the relationship between the speed of spending and administrative capacity. More in detail, previous studies have highlighted a positive correlation between administrative capacity and financial absorption (Milio 2007; Terracciano & Graziano 2016; Hagemann 2019; Incaltarau et al. 2020; Surubaru 2017; Tignasu 2018), ignoring the impact it may have on spending speed. Lastly, our results appear innovative for the question of whether the speed of public spending varies at different levels of government with a certain quality of institutions and administrative capacity. Some articles have examined the efficiency of spending at a specific government level. For example, Belanguer-Coll et al. (2010) analysed the efficiency of Spanish municipalities, but did not compare municipalities with other levels of government. Our findings have implications for the debate on the absorption of EU resources, both from cohesion policy and Next Generation EU recovery plan. The amount of EU funds is growing over time, but many countries still face considerable difficulties in spending these resources, especially in the early years of the programming periods. Italy has a problem with late spending, which reduces the effectiveness of EU funds. Decentralization may help to explain the late spending problem in some countries and regions characterized by low quality of local governments and scarce administrative capacity.

The remainder of the paper is structured as follows. Section 2 discusses the literature on the role of administrative capacity, institutional quality and government decentralization on public spending. Section 3 describes the dataset and the methodology, and Section 4 provides the main results. Section 5 concludes and draws some policy implications.

2. Literature review

2.1. Quality of institutions and public spending

A strand of literature has shown that local institutional quality affects economic growth through its outcome on different policies and investments, such as interventions to promote innovative and entrepreneurial behaviour (Baumol 1990), entrepreneurship (Nistotskaya et al. 2015; Aparicio et al. 2016; Huggins & Thompson 2016), innovation (Rodríguez-Pose & Di Cataldo 2015), regional competitiveness (Annoni and Dijkstra, 2017), productivity (Kaasa 2016), industrial diversification (Cortinovis et al. 2017), resilience (Ezcurra & Rios 2019), or infrastructure (Crescenzi et al. 2016). Corradini (2021) finds robust evidence of a causal effect of formal institutional quality over economic growth, across a panel of Italian NUTS-3 regions, supporting the hypothesis that 'institutions rule' also at the sub-national level. This effect is found to be more pronounced in areas with lower levels of economic development and no evidence is found of a reverse effect of growth on local institutional quality. In addition, several contributions highlight that local government quality is a fundamental driver of economic growth (Ketterer & Rodríguez-Pose 2018) and that the connection between the quality of local institutions and economic performance is achieved through how variations in government quality shape the design, implementation, and monitoring of public policies.

Over the last two decades, there has been a growing focus on the role of government quality as a prerequisite for the effectiveness of EU Cohesion Policy, besides being a determinant of economic growth. Several studies have documented a relationship between the quality of government and the absorption of European funds, both at the national level and in some case study regions (Baun & Marek 2017; Milio 2007; Terracciano & Graziano 2016). Institutional quality and good governance are important factors in explaining how weak regions can have momentum through spending public funds (Crescenzi et al. 2020). Rodríguez-Posé & Garcilazo (2015), for example, demonstrated that poor institutional quality can imply the ineffectiveness of EU regional development funds, and Accetturo et al. (2014) suggested that the quality of local government affects the economic impact of EU Structural Funds. Empirical evidence has shown that regions with better quality institutions obtain larger funding allocations of European Structural and Investment Funds (Bouvet & Dall'Erba 2010; Charron, 2016), have higher levels of absorption (Terracciano & Graziano 2016; Tosun 2014), and realise faster economic growth (Bachtrögler 2016; Becker et al. 2013; Rodríguez-Pose & Garcilazo 2015). Del Monte et al. (2022 a) demonstrated that the quality of governance contributes to explaining the different paces of spending across Italian regions. In Southern Italy, the low quality of institutions was associated with longer project durations, thus indicating less efficient local governments.

2.2. Administrative capacity and public spending

A growing number of studies are focusing on the impact of administrative capacity on public funding. Bachtler et al. (2023) focus on the internal and external administrative factors that contribute differently to measures of performance – absorption, regularity and effectiveness- of EU Cohesion Policy. The authors argue that the internal components of administrative capacity are variously defined and recognised as human, organisational structures, physical resources, systems and tools, leadership, and openness to external knowledge. Other factors influencing the ability of the system to achieve its objectives have been identified as external to the administrative system and are the quality of public administration, legal stability/regulatory quality, centralisation/decentralisation of governance, political influence.

A considerable portion of the research conducted thus far on Cohesion Policy has concentrated on examining the significance of administrative capacity in shaping the outcomes of European Structural and Investment Funds (ESIF) programs. These studies assert that the success of the policy is conditional on the ability of national, regional and local administrations to design robust strategies, allocate resources effectively, administer EU funding efficiently and ensure better financial compliance (e.g., Bachtler et al. 2014; Mendez & Bachtler 2017; Milio 2007; Polverari et al. 2020; Surubaru 2017; Terracciano & Graziano 2016; Tosun 2014). Incaltarau et al. (2020) study the influence of administrative capacity and political governance on the absorption of structural and cohesion funds, recommending a focus on capacity-building and anti-corruption efforts. Bachtrögler-Unger et al. (2022) show that the regional administrative capacity is significant as a mediator between regional growth and the effect of Cohesion Policy on growth and is particularly important when regions implement measures to support local manufacturing firms, as these measures are more effective in a better institutional environment. Aivazidou et al. (2020) found that low administrative capacity of regional strategies 'may increase the absorption rate, though without supporting regional growth'. Some authors indicate a favourable correlation between administrative capability and financial absorption, as demonstrated by analyses conducted in two Italian regions (Milio 2007;

Terracciano & Graziano 2016), in Eastern European member states (Surubaru 2017; Tignasu 2018) and Central European member states (Hagemann 2019). Mendez & Bachtler (2022) analyses the relationship between the quality of government and multiple dimensions of administrative performance in Cohesion Policy. Their results show that government quality is a key determinant of administrative performance in terms of financial compliance, timely spending and outcomes.

2.3. Decentralization and public spending

The assertion that decentralization brings about remarkable economic advantages is founded in the classical fiscal federalism literature. In 1959, Musgrave proposed a comprehensive theory of State and public finance, delineating three primary functions of government: the provision of public goods and services (resource allocation), income redistribution, and the stabilization of economic activities to mitigate fluctuations in the business cycle (macroeconomic stabilization). According to this theory, the central government ought to handle income redistribution and macroeconomic stabilization, while revenue assignment should be decentralized to other levels of government.

The contributions of Tiebout (1956) and Oates (1972) emphasized that local governments can align public spending more effectively with the preferences of individuals residing in diverse territories. Decentralization, furthermore, purportedly enhances the efficiency in mobilizing underutilized resources through competition among subnational governments (Oates, 1996; Rodríguez-Pose & Ezcurra 2010). Some authors have argued that decentralization leads to greater dynamism and well-being for citizens wherever they live (Oates 1972; Putnam 1993). In a decentralized system, local administrations can better identify the needs and preferences of citizens and offer public services that are more suited to the needs of the local community. Citizens, moreover, thanks to a greater "proximity", can better monitor and evaluate the services managed and offered by local levels of government. This in turn gives governments a greater incentive to operate efficiently. Intensified competition among subnational entities can also stimulate productive efficiency as regions specialize in their comparative advantages (Rodríguez-Pose & Bwire 2004). Faguet (2004) shows that decentralization led to higher investment in human capital and social services as the poorest regions of Bolivia chose projects according to their greatest needs. Additionally, the heightened proximity of political power to citizens may enhance accountability and transparency, foster increased participation, and mitigate corruption (Ebez & Yilmaz 2002; Putnam 1993).

Another strand of literature emphasizes that the effect of decentralization can be negative, especially in large countries (Rodriguez-Pose & Gill 2005). Local governments may face challenges in efficiently managing complex public services due to financial limitations, technical constraints, or a deficit of qualified personnel. Conversely, central governments can achieve substantial economies of

scale in providing public services, often offering them at a lower cost compared to local levels of government. Viesti (2019) highlights that the devolution of functions and resources to subnational governments can also undermine the redistributive power of the State.
Transparency and the reduction of corruption may not necessarily improve through decentralization, particularly when resources are limited. Inadequate human and financial capacity can hinder efforts to effectively monitor and prevent corruption (De Mello & Barenstein 2001). Del Monte and Papagni (2007) show that the experience of the Italian regions since their establishment in the 1970s has not been positive, with a growing trend towards corruption. Corruption is often linked to the closeness between politicians, bureaucrats, and citizens. The contiguity in local communities means that individuals or organized groups can "capture" politicians and bureaucrats for their interests.
A further argument supporting the negative view of decentralization is that national administrations often offer better opportunities than local administrations and attract more qualified and skilled workers (Prud'homme 1995). In a decentralized system, the competition among regions can disadvantage weaker areas while favouring wealthier ones, thereby exacerbating regional disparities. This phenomenon is especially pronounced in countries like Italy, where there are significant disparities between the North and South, with the latter experiencing considerable lag in development. Charron (2014) studied the quality of local administrations in more than 200 European regions, and found that the top Italian region was in 118th place. Southern Italian regions were at the lowest levels of this ranking. The recent experience of regional policies during the COVID-19 pandemic has also shown a lot of inefficiencies, both in Italy and in other countries such as the United States (Agnew 2022). Tanzi (1996; 2001) pointed out that the disadvantages of decentralization exceed the advantages if the quality of local institutions is low. Sacchi et al. (2019) corroborated this view by analysing the relationship between decentralization and the provision of local public services when the quality of institutions varied. They found that when the institutional context was unfavourable—that is, there was high corruption, reduced independence of mass media, low quality of politicians, and low political participation of citizens— decentralization led to a reduction in the quantity and quality of local public services provided to citizens. Rodriguez-Pose and Mustra (2022) emphasize that the economic benefits of decentralization are more pronounced in areas with highly efficient local governments. They further suggest that when surrounded by other regions characterized by both high levels of decentralization and high government quality, the positive impacts of transferring powers and resources to subnational tiers of government are amplified. This phenomenon can potentially surpass the effects of directly transferring authority and resources to local governments, as well as the quality of those local governments themselves.

Finally, Rodriguez-Pose and Vidal-Bover (2022) showed that the disadvantages overcame the advantages if the State did not provide local governments with adequate monetary resources to manage the functions that were delegated to them. This negative effect was stronger in politically and economically decentralized regions and in regions with the greatest wealth effect.
However, one key factor behind decentralized governments has generally been overlooked in past scholarly literature: whether the efficiency of public spending depends on the government decentralized and what is the role of the quality of the institution. Del Monte et al. (2022 a) investigate whether government decentralization affects the duration of public projects in Italy. They find regional differences in project duration and these regional differences were driven by the quality of local institutions.

3. Data and methodology

To evaluate the speed of spending of the various levels of government, a rich dataset from a variety of sources is constructed. Data on the speed of spending come from OpenCoesione,[5] the open government initiative on EU cohesion policies in Italy coordinated by the Department for Cohesion Policy of the Presidency of the Council of Ministers. The data are available for all programming cycles from 2007 onward (to 02/2023 for this paper) and the database is updated bimonthly. This dataset encompasses a wide range of information related to public projects managed by different levels of administration, such as regional, provincial, and municipal. Furthermore, this source provided information on the type of funding (ERDF, FSC, ESF, National Fund), the amount of funding, the level of expenditure, the location of the project, and the type of managing organization. Extensive information is also provided on the timing of implementation and payments for individual projects. Further information can be obtained on the sector in which the project is classified and the thematic area of Cohesion Policy expenditure. In our analysis, we focus on the public projects financed under the 2007-2013 EU programming period since many projects from the subsequent programming period (2014-2020) are still ongoing. A total of 59,236 projects are included in our sample.

3.1 Dependent variable and regressors

The dependent variable focuses on the duration of public projects and their association with various levels of government overseeing them. More in detail, the dependent variable measures project duration in terms of the number of days between the start and the end of the projects (variable *Duration*).

[5] https://opencoesione.gov.it/it/

The choice of the main explanatory variables is focused on capturing three key dimensions: levels of government, the quality of institutions, and administrative capacity. Binary variables indicating which level of government is managing each project are included in the model to measure the effect of decentralization. The different levels of government considered include regional governments (*Region*), provincial governments (*Province*), and municipal governments (*Municipality*).

To capture the quality of institutions we use the composite indicator proposed by Nifo and Vecchione (2014), which assesses the quality of local institutions in Italy at subnational levels (variable *IQI*). This indicator is similar to the World Governance Indicators (WGI) but focuses specifically on Italy and uses actual data rather than survey-based assessments. The indicator comprises five dimensions: 1) Participation in public elections and cultural vitality; 2) Quality of public services in health, waste management, and environment; 3) Government's ability to promote policies benefiting firms and the private sector; 4) Rule of law, including crime rates, tax evasion, and the shadow economy; 5) Degree of corruption among public officials.

The last factor of main interest is the administrative capacity of the different levels of government. Administrative capacity is essential for the successful implementation of government policies. A rapid implementation of public projects is considered crucial to achieving predetermined objectives, ensuring a positive impact on the communities involved. Effective administrative competence is equally important, as it contributes to managing projects efficiently, reducing delays, and maximising the use of available resources. The synergy between the speed of spending and a robust administrative capacity is essential to ensure that public investments translate successfully into tangible results. Ultimately, efficient management of public funds requires a balance between speed and capacity, ensuring that resources are used responsibly and effectively for the benefit of the community.

Administrative capacity is a multidimensional concept captured in this analysis by using a range of variables built from multiple sources. To proxy administrative capacity, we include in the model variables measured at the level of individual public administrations. The data for these variables are from two key sources: the annual personnel accounts of public administrations, processed by the State General Accounting Office (SGAO), and datasets from the Italian National Statistical Institute (ISTAT). Information from both sources is compiled for each public entity, covering the period from 2010 to 2019. The variables include the percentage of employees holding a university degree (*Education*), the percentage change in the number of public employees during project implementation (*Δ employees*), the total number of employees within the administration (*Total employees*), the average age of public employees (*Average age*), and the percentage of employees with more than 20 years of service (*Seniority*). The analysis aims to investigate how these factors influence the duration of public projects, providing insights to improve the efficiency and effectiveness of European fund utilization.

The regression model includes a comprehensive set of controls to account for various characteristics of the projects under scrutiny. These controls were essential for understanding the nuances and complexities inherent in the projects and their contexts. Here's a breakdown of the key components integrated into the model. Firstly, the variable *Amount* was utilized as a proxy for project size, representing the total financing allocated to each project, encompassing both public and private funds. This variable provided insights into the scale and scope of the projects under examination. Secondly, the variable *Beneficiaries* was employed to capture the number of beneficiaries associated with each project. This variable was used to understand which subjects are responsible for the start-up, implementation and functionality of the project. Thirdly, we control for the share of private funds on the total amount of the project (*Private*). Projects involving a large share of private funds are expected to be managed more efficiently and therefore to have a lower duration.

Additionally, projects were categorized into 13 thematic areas, such as digital agenda, environment, culture and tourism, among others. Binary indicators for these thematic areas are included in the model. This categorization allows for the differentiation of projects based on their thematic focus and provides insights into the varying degrees of complexity and resource requirements across different project types. Moreover, the model accounts for the sources of financing of the projects, distinguishing between EU structural funds, cohesion funds, and national funds. Binary indicators are utilized to control for these categories of financing sources, enabling the analysis to capture the influence of funding mechanisms on project outcomes and implementation dynamics. Table 1 provides summary statistics and correlations.

Table 1. Descriptive statistics and correlations

	Mean	St. dev.	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Duration	454	481	10	7356											
(2) Region	0.48	0.49	0	1	0.04										
(3) Province	0.35	0.47	0	1	-0.15	-0.72									
(4) IQI	0.60	0.21	0	1	-0.07	-0.27	0.43								
(5) Education	0.44	0.11	0	1	0.07	-0.18	-0.05	-0.21							
(6) Average age	49.4	3.9	25	62	-0.01	0.07	-0.20	-0.40	0.32						
(7) Seniority	21.4	6.5	1	40	-0.04	0.01	0.14	0.16	-0.17	0.09					
(8) Δ employees	-2.10	17.8	-93	770	-0.08	0.20	-0.18	-0.01	-0.09	-0.06	-0.04				
(9) Total employees	1440	2425	10	17201	0.01	0.31	-0.16	-0.01	-0.08	0.02	-0.15	0.04			
(10) Amount	84	1209	12	214000	0.05	-0.01	-0.03	-0.04	0.03	0.01	-0.03	-0.01	0.04		
(11) Private	0.01	0.04	0	0.9	-0.01	0.03	-0.06	-0.08	0.06	0.04	-0.02	0.01	0.06	0.01	
(12) Beneficiaries	4.9	2.12	0	14	0.02	-0.04	0.02	0.01	0.04	-0.03	-0.04	-0.01	-0.04	0.01	0.07

3.2. Methodology

We conducted an econometric analysis to ascertain the impact of administrative capacity, institutional quality and government decentralization on the duration of public projects. As explained above, we have a cross-section of public projects implemented in the period 2010-2019. The observations of the model are therefore the public projects, and the dependent variable measures their duration. We use ordinary least squares (OLS) to estimate the model, with standard errors clustered at the municipal level. To reduce the problems of unobserved heterogeneity and omitted variables, the model included fixed effects at the provincial level, which account for the variations in the local contexts and environments that could potentially impact project outcomes, as well as dummies for the types of projects and sources of funding.

The model is estimated in various steps (Tab.2). In the first step, we estimate the impact of decentralization and the quality of institutions on project duration (column 1). In the second step, we replace the quality of institutions with the variables capturing administrative capacity (column 2), and in the third step, we jointly consider the quality of institutions and administrative capacity (column 3). Finally, the model is estimated for each of the three levels of government (municipalities, provinces, and regions), as shown in Table 3.

4. Results

Table 2 illustrates the estimates for the full sample including all public projects managed by municipal, regional, and provincial governments. The dummies *Region* and *Province* are statistically significant with a negative sign in all regressions, meaning that projects managed by regional and provincial governments have a shorter duration than projects managed by municipal governments. In each regression, in fact, the category of managing institution excluded concerns municipalities, thus the signs associated with *Region* and *Province* must be interpreted against this level of government. Therefore, projects conducted in regions and provinces are completed more quickly than those carried out in municipalities. The estimated coefficients suggest also that regional governments are the most efficient in managing public projects, while municipal governments are the least efficient. Government decentralisation therefore is associated with a lower speed of public spending. Another noteworthy finding is the influence of the quality of local institutions on project duration. In Column (1) the quality of institutions, measured by *IQI*, has a negative and significant impact on project duration. This suggests enhanced efficiency in project management as the quality of institutions increases. Consequently, governmental efforts aimed at enhancing project efficiency should prioritize improvements in the quality of local institutions rather than decentralizing functions. In Column (2) we introduce several proxies of administrative capacity that are statistically significant and contribute to reducing the duration of public projects. A high percentage of graduate workers is associated with the few days required to complete a project, suggesting that human capital may accelerate project implementation (*Education*). Furthermore, a higher percentage change of employees during a project development (*Δ employees*) is associated with a reduction in completion time, indicating that the stock of employees is important to reduce the duration of public projects. The age of the employees (*Average age*) also plays a relevant role: a higher average age is correlated with a lower number of days required to complete a project, suggesting that experience and expertise may accelerate the process. Furthermore, the analysis suggests that the seniority of employees (*Seniority*) may influence the completion time, with a higher percentage of employees with long seniority correlated with a lower number of days required to complete a project. This again suggests a positive contribution of accumulated experience over time. Lastly, the total number of workers is used as a proxy for the size of governments; the variable is not statistically significant (*Total employees*). These results are generally confirmed in Column (3) when we consider together the quality of institutions and administrative capacity.

It is interesting to observe that in Column (3), when we include together all regressors, *IQI* and all proxies for administrative capacity keep their statistical significance and negative sign. The coefficient associated with *IQI* is now lower than in Column (1), while the coefficients associated with the proxies of administrative capacity do not change. These findings suggest that the quality of institutions and administrative capacity are distinct drivers of project duration, and therefore of spending speed; in addition, administrative capacity mediates the impact of institutional quality.

As for the controls, *Amount* and *Beneficiaries* have positive and significant coefficients, indicating that project duration increased with increased total funding and number of beneficiaries. The share of private funding instead negatively impacts project duration, meaning that the involvement of private subjects in a project reduces its duration.

Table 2. Determinants of project duration

	(1)	(2)	(3)
Region	–148.7***	–179.6***	–182.0***
	(33.23)	(33.27)	(32.73)
Province	–145.7***	–171.6***	–172.8***
	(25.45)	(28.75)	(28.76)
IQI	–139.2**	-	–84.99*
	(62.85)		(48.07)
Education	-	–2.751***	–2.608***
		(0.788)	(0.770)
Average age	-	–110.5***	–109.89***
		(27.58)	(49.82)
Seniority		–89.04*	–91.36*
		(46.65)	(49.83)
Δ employees	-	–1.625**	–1.627**
		(0.708)	(0.709)
Total employees		0.003	
		(0.003)	
Amount	0.014*	0.013*	0.013*
	(0.008)	(0.006)	(0.006)
Private	–175.7*	–170.9	–174.9*
	(107.1)	(111.0)	(108.1)
Beneficiaries	98.29***	66.81*	66.65**
	(27.84)	(35.58)	(33.41)
Dummies for the types of projects	Yes	Yes	Yes
Dummies for the sources of funding	Yes	Yes	Yes
Provincial dummies (NUTS 3 level)	Yes	Yes	Yes
Observations	59,236	59,236	59,236
R^2	0.16	0.18	0.19

Notes: The dependent variable measures the duration of the projects in days. Standard errors clustered at the municipality level in parentheses. ***, **, * statistically significant at the 1%, 5% and 10% level.

Table 3 shows the second set of estimates, where separate regressions are presented for the different levels of government: municipalities, provinces, and regions. For each level of government, the estimations are structured in two steps. In the first step, only the quality of institutions is considered through the variable *IQI*; in the second step the proxies for administrative capacity are added in the regression model.

IQI consistently exhibits significant and negative coefficients across all regressions, suggesting that the quality of institutions is important in reducing project durations for all levels of government. The impact varies considerably across different administrative levels, being stronger for municipal governments and weaker for regional governments.

When the measures of administrative capacity are included in the regressions, the coefficient of *IQI* is lower and it is no longer significant for provincial and regional governments. These findings confirm that administrative capacity plays a mediating role in the relationship between the quality of institutions and project duration.

The variables for administrative capacity in general have negative and statistically significant coefficients, confirming the important role in reducing the duration of public projects. The coefficient for *Education* is negative and statistically significant, suggesting that a higher percentage of graduate workers is associated with shorter project durations; this is more evident for regional governments (–17.50 days) than in provincial governments (–10.70) and municipal governments (–3.042). Similarly, *Δ employees* is significant in all regressions and the impact on project duration is higher for provincial and regional governments.

Average age significantly reduces project duration only in the regression for provincial governments and *Seniority* reduces project duration only for provincial and regional governments. A high percentage of employees with over 20 years of seniority is linked to shorter project durations, but this trend is observed only in provinces and regions. Municipalities, on the other hand, do not show statistically significant results in this regard. This highlights potential differences in workforce dynamics and project management practices across administrative levels. *Total employees* do not show a statistically significant coefficient, suggesting that the total number of employees does not significantly influence project durations.

In sum, these findings suggest that institutional quality and administrative capacity have an important role in reducing the duration of public projects and therefore in increasing the speed of spending of public funds. The impact of the quality of institutions appears stronger at the lowest levels of government, while administrative capacity seems more effective at the highest levels of government.

Table 3. Determinants of project duration for the different levels of government

	Municipalities		Provinces		Regions	
IQI	–1697.2***	–1458.9***	–429.1***	–36.86	–115.7*	–76.42
	(497.1)	(412.1)	(209.0)	(247.1)	(66.88)	(193.1)
Education	-	–3.042***	-	–10.70***	-	–17.50**
		(0.578)		(4.221)		(8.464)
Average age	-	81.70	-	–251.7*	-	34.27
		(67.11)		(149.7)		(28.60)
Seniority	-	10.91	-	–695.0***	-	–1192.5***
		(37.22)		(233.1)		(197.3)
Δ employees	-	–1.543**	-	–14.61***	-	–1.874***
		(0.643)		(3.975)		(0.333)
Total employees	-	0.001	-	0.690	-	0.017
		(0.001)		(0.648)		(0.011)
Amount	0.210***	0.202***	0.197***	0.153***	0.004**	0.001
	(0.039)	(0.038)	(0.033)	(0.032)	(0.002)	(0.001)
Private	–165.7	–191.8	158.7	–154.6	100.7	293.0**
	(232.6)	(217.6)	(294.0)	(283.1)	(109.2)	(137.0)
Beneficiaries	–41.16	–50.78	16.51	–42.70	114.6	166.2***
	(56.38)	(54.02)	(91.80)	(92.54)	(15.8)	(24.68)
Dummies for the types of projects	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for the sources of funding	Yes	Yes	Yes	Yes	Yes	Yes
Provincial dummies (NUTS 3 level)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,198	9,198	21,231	21,231	28,807	28,807
R^2	0.31	0.34	0.36	0.55	0.13	0.16

Notes: The dependent variable measures the duration of the projects in days. Standard errors clustered at the municipality level in parentheses. ***, **, * statistically significant at the 1%, 5% and 10% level.

5. Conclusions

This paper has analyzed the impact of internal and external administrative factors on the speed of public spending for the various levels of Italian government (municipalities, provinces, and regions). The speed of spending is computed for projects under OpenCoesione, the open government initiative focused on EU cohesion policies in Italy. The analysis encompasses projects financed during the 2007-2013 EU programming period, in the period 2010-2019.

Overall, external factors captured by the Institutional Quality Index, which captures aspects such as public services, local economic activity, justice, corruption, cultural level, and citizen participation in public life, exert a negative and significant impact on project duration. Therefore, an improvement in institutional quality tends to decrease the speed of public spending. Consequently, governmental efforts aimed at enhancing project efficiency should prioritize improvements in the quality of local institutions rather than decentralizing functions. This is particularly true for the lowest levels of government, like municipal governments. In addition to external factors, several internal administrative factors also influence the speed of public spending. A higher percentage of graduate workers correlates with shorter project durations. Moreover, the average age of employees indicates that as it increases, project durations tend to increase as well. Additionally, factors such as seniority of personnel and percentage change in employees exhibit significant impacts on project duration. These results underscore the significant role played by the characteristics of employed personnel in determining the overall duration of public projects. Taken together these findings indicate that administrative capacity is an important factor in reducing project duration, and therefore in enhancing the speed of public spending. The impact of administrative capacity is particularly strong at the higher levels of government, like provincial and regional governments.

The implications of our results extend beyond academic discourse to inform concrete policy actions. It is imperative for both local and central administrations to proactively implement targeted policies aimed at bolstering institutional quality and enhancing administrative capacity. By prioritizing initiatives that foster transparency, accountability, and efficiency within administrative structures, governments can expedite project implementation and optimize the utilization of public resources for the benefit of society at large.

The analysis had some limitations. First, our analysis may be constrained by data availability and quality, particularly regarding certain internal administrative factors such as workforce demographics. Future research could benefit from more comprehensive and detailed datasets to provide a more nuanced understanding of these dynamics. Second, our study focuses on a specific type of fund and period, which may limit the generalizability of our findings. Future research could explore how the influence of administrative factors varies across different funds, such as the National

Program for Recovery and Resilience. Third, our analysis may not capture all relevant administrative factors that influence public spending dynamics. Future research could explore the impact of additional variables, such as bureaucratic procedures, project management techniques, or political factors, to provide a more comprehensive understanding of the administrative determinants of project duration. Finally, while we highlight the importance of improving institutional quality and administrative capacity, the effectiveness of specific policy interventions in achieving these objectives remains uncertain. Future research could evaluate the impact of various policy initiatives aimed at enhancing administrative performance and project efficiency. Addressing these limitations through further empirical research and methodological advancements can deepen our understanding of the complex interplay between administrative factors and public spending dynamics, ultimately informing more effective policy interventions and administrative practices.

References

Accetturo, A., De Blasio, G., & Ricci, L. (2014). A tale of an unwanted outcome: Transfers and local endowments of trust and cooperation. *Journal of Economic Behavior & Organizati*on, 102, 74-89.

Agenzia per la Coesione Territoriale (2018). *Rapporto sui tempi di attuazione delle opere pubbliche*. Temi CPT n. 6.

Agnew, J. (2022). Failing federalism? US dualist federalism and the 2020–22 pandemic. *Regional Studies Regional Science*, 9(1), 149–171.

Aivazidou, E., Cunico, G., & Mollona, E. (2020). Beyond the EU structural funds' absorption rate: How do regions really perform?. *Economies*, 8(3), 55-75.

Annoni, P. and Dijkstra, L. (2017). Measuring and Monitoring Regional Competitiveness in the European Union. In Huggins, R. and Thompson, P. (eds) Handbook of Regions and Competitiveness (Cheltenham: Edward Elgar Publishing), 49–79.

Aparicio, S., Urbano, D. & Audretsch, D. (2016). Institutional Factors, Opportunity Entrepreneurship and Economic Growth: Panel Data Evidence. *Technological Forecasting and Social Change*, 102, 45–61.

Bachtler, J., Ferry, M., & Gal, F. (2018). Financial implementation of European Structural and investment funds. *European Parliament, Policy Department for Structural and Cohesion Policies.* 10.2861/55

Bachtler, J., Mendez, C., & Oraže, H. (2014). From conditionality to Europeanization in central and Eastern Europe: Administrative performance and capacity in Cohesion Policy. *European Planning Studies*, 22(4), 735–757. https://doi.org/10. 1080/09654313.2013.772744

Bachtler, J., Polverari, L., Domorenok, E., & Graziano, P. (2023). Administrative capacity and EU Cohesion Policy: implementation performance and effectiveness. *Regional Studies*, 1-5.

Bachtrögler, J. (2016). On the effectiveness of EU structural funds during the Great Recession: Estimates from a heterogeneous local average treatment effects framework (Working Paper No. 230).

Bachtrögler-Unger, J., Fratesi, U., & Perucca, G. (2022). Administrative capacity and the territorial effects of EU support to firms: A two-step analysis. Regional Studies.

Baun, M., & Marek, D. (2017). The limits of regionalization: The intergovernmental struggle over EU Cohesion Policy in the Czech republic. *East European Politics and Societies: And Cultures*, 31(4), 863–884. doi:10.1177/0888325417720717

Balaguer-Coll, M. T., Prior, D., & Tortosa-Ausina, E. (2010). Decentralization and efficiency of local government. *The Annals of Regional Science,* 45, 571-601.

Baun, M., & Marek, D. (2017). The limits of regionalization: The intergovernmental struggle over EU Cohesion Policy in the Czech Republic. *East European Politics and Societies,* 31(4), 863-884.

Baumol WJ (1990) Entrepreneurship: productive, unproductive, and destructive. *Journal of Political Economy* 98: 893–921

Becker, S. O., Egger, P. H., & von Ehrlich, M. (2013). Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy*, 5(4), 29–77. https://doi.org/10.1257/pol.5.4.29

Bouvet, F., & Dall'Erba, S. (2010). European regional structural funds: How large is the influence of politics on the allocation process? *Journal of Common Market Studies*, 48(3), 501–

528. https://doi.org/10.1111/j.1468-5965.2010.02062.x

Charron, N. (2016). Explaining the allocation of regional structural funds: The conditional effect of governance and self-rule. *European Union Politics*, 17(4), 638–659. https://doi.org/10.1177/1465116516658135

Charron, N., Dijkstra, L., & Lapuente, V. (2014). Regional governance matters: Quality of government within European Union member states. *Regional studies*, 48(1), 68-90.

Corradini, C., Santini, E., & Vecciolini, C. (2021). The geography of Industry 4.0 technologies across European regions. *Regional Studies,* 55(10-11), 1667-1680.

Cortinovis, N., Xiao, J., Boschma, R. & van Oort, F.G. (2017). Quality of Government and Social Capital as Drivers of Regional Diversification in Europe. *Journal of Economic Geography*, 17(6), 1179–208.

Crescenzi, R., Di Cataldo, M. & Rodríguez-Pose, A. (2016). Government Quality and the Economic Returns of Transport Infrastructure Investment in the European Union. *Journal of Regional Science*, 56, 555–582.

Crescenzi, R., Fratesi, U., & Monastiriotis, V. (2020). Back to the member states? Cohesion Policy and the national challenges to the European Union. *Regional Studies*, *54*(1), 5-9.

De Mello, L., & Barenstein, M. (2001). Fiscal decentralization and governance: A cross-country analysis (No. 2001-2071). Washington, DC: International Monetary Fund.

Del Monte, A., De Iudicibus, A., Moccia, S. & Pennacchio, L. (2022a). Speed of spending and government decentralization: evidence from Italy. *Regional Studies,* 56(12), 2133-2146.

Del Monte, A., De Iudicibus, A., Moccia, S., & Pennacchio, L. (2022b). Decentralization and Duration of Public Projects in Italy. *Rivista economica del Mezzogiorno,* 36(4), 975-992.

Del Monte, A., & Papagni, E. (2007). The determinants of corruption in Italy: Regional panel data analysis. *European Journal of Political Economy*, 23(2), 379–396.

Ebel R D and Yilmaz S (2002) On the Measurement and Impact of Fiscal Decentralization. *Policy Research Working Paper Series* 2809: 1‑26

Ezcurra, R., & Rios, V. (2019). Quality of Government and Regional Resilience in the European Union. Evidence from the Great Recession. *Papers in Regional Science*, 98(3), 1267–90.

Faguet, J.-P. (2004). Does decentralization increase government responsiveness to local needs?: Evidence from Bolivia. *Journal of Public Economics*, 88(3–4), 867–893.

Hagemann, C. (2019). How politics matters for EU funds' absorption problems–A fuzzy-set analysis. *Journal of European Public Policy*, 26(2), 188-206.

Huggins, R., & Thompson, P. (2016). Socio-spatial culture and entrepreneurship: some theoretical and empirical observations. *Economic Geography,* 92(3), 269-300.

Incaltarau, C., Pascariu, G. C., & Surubaru, N. C. (2020). Evaluating the determinants of EU funds absorption across old and new member states–The role of administrative capacity and political governance. *Journal of Common Market Studies*, *58*(4), 941-961.

Kaasa, A. (2016). Social Capital, Institutional Quality and Productivity: Evidence from European Regions. *Economics & Sociology*, 9(4), 11–26.

Ketterer, T.D., & Rodríguez-Pose, A. (2018). Institutions vs. 'First-Nature' Geography: What Drives Economic Growth in Europe's Regions?. *Papers in Regional Science*, 97, S25–S62.

Mendez, C., & Bachtler, J. (2017). Financial Compliance in the European Union: A Cross-National Assessment of Financial Correction Patterns and Causes in Cohesion Policy. *Journal of Common Market Studies*, 55(3), 569-592.

Mendez, C., & Bachtler, J. (2022). The quality of government and administrative performance: explaining Cohesion Policy compliance, absorption and achievements across EU regions. *Regional Studies*, *58*(4), 690–703. https://doi.org/10.1080/00343404.2022.2083593

Milio, S. (2007). Can administrative capacity explain differences in regional performances? Evidence from structural funds implementation in southern Italy. *Regional studies*, 41(4), 429-442.

Musgrave R. (1959), The Theory of Public Finance. New York: McGraw Hill.

Nifo, A., & Vecchione, G. (2014). Do institutions play a role in skilled migration? The case of Italy. *Regional Studies*, 48(10), 1628–1649.

Nistotskaya, M., Charron, N. & Lapuente, V. (2015). TheWealth of Regions: Quality of Government and SMEs in 172 European Regions. *Environment and Planning C: Government and Policy,* 33(5), 1125–1155.

Oates, W. E. (1972). Fiscal federalism. Harcourt, Brace, Jovanovich. Palermo, F., & Wilson, A. (2013). The dynamics of decentralization in Italy: Towards a federal solution? *(European Diversity and Autonomy Papers Series,* WP No. 4).

Oates, W. E. (2004). Estimating the demand for public goods: The collective choice and contingent valuation approaches. In *Environmental Policy and Fiscal Federalism* (pp. 78-99). Edward Elgar Publishing.

Polverari, L. (2020). The multiple accountabilities of EU cohesion policy: Wicked problem or north star?. *In Financial Accountability in the European Union* (pp. 111-126). Routledge.

Prud'homme R. (1995). The dangers of decentralization. *The World Bank Research Observer*, 10(2), 201–220.

Putnam, R. D., Leonardi, R., & Nanetti, R. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton University Press.

Rodríguez-Pose, A., & Bwire, A. (2004). The economic (in) efficiency of devolution. *Environment and Planning a*, 36(11), 1907-1928.

Rodriguez-Posé, A., & Ezcurra, R. (2010). Does decentralization Matter for regional disparities? A cross-country analysis. *Journal of Economic Geography*, 10(5), 619–644. https://doi.org/10.1093/jeg/lbp049

Rodríguez-Pose, A., & Garcilazo, E. (2015). Quality of government and the returns of investment: Examining the impact of cohesion expenditure in European regions. *Regional Studies*, 49(8), 1274–1290. https://doi.org/10.1080/00343404.2015.1007933

Rodríguez-Pose, A. & Di Cataldo, M. (2015). Quality of Government and Innovative Performance in the Regions of Europe. *Journal of Economic Geography*, 15(4), 673–706.

Rodríguez-Posé, A., & Gill, N. (2005). On the 'economic dividend' of devolution. *Regional Studies,* 39(4), 405–420. https://doi.org/10.1080/00343400500128390

Rodríguez-Pose, A., & Muštra, V. (2022). The economic returns of decentralisation: Government quality and the role of space. *Environment and Planning A: Economy and Space*, 54(8), 1604-1622.

Rodríguez-Pose, A., & Vidal-Bover, M. (2022). Unfunded Mandates and the Economic Impact of Decentralisation. When Finance Does Not Follow Function. *Political Studies*.

Sacchi, A., Scalera, D., & Zazzaro, A. (2019). Decentramento, qualità istituzionale e fornitura di servizi pubblici: Alcune evidenze empiriche. *Rivista Economica del Mezzogiorno*, 33(1), 81–104.

Surubaru, N. C. (2017). Administrative capacity or quality of political governance? EU Cohesion Policy in the new Europe, 2007–13. *Regional Studies*, 51(6), 844-856.

Tanzi V. (1996). Fiscal Federalism and Decentralization: A Review of Some Efficiency and Macroecomic Aspects. *The International Bank for Reconstruction and Development, The World Bank*, *295–315*.

Tanzi, V. (2001). Pitfalls on the road to fiscal decentralization. Economic Reform Project Working Paper no. 19, Carnegie Endowment for International Peace.

Terracciano, B., & Graziano, P. R. (2016). EU Cohesion Policy implementation and administrative capacities: Insights from Italian regions. *Regional & Federal Studies*, 26(3), 293–320. https://doi.org/10.1080/13597566.2016.1200033

Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5), 416–424. https://doi.org/10.1086/257839

Tiganasu, R., Incaltarau, C., & Pascariu, G. C. (2018). Administrative capacity, structural funds absorption and development. Evidence from central and eastern European countries. Romanian J. Eur. Aff., 18, 39.

Tosun, J. (2014). Absorption of regional funds: A comparative analysis. *Journal of Common Market Studies*, 52(2), 371–387. https://doi.org/10.1111/jcms.12088

Viesti G. (2019). Verso la secessione dei ricchi? Autonomie regionali e unità nazionale, Roma-Bari, Laterza.

APPENDIX

Description	Variable name	Source
Personnel feminization rate	FEMPER	ISTAT (Italian National Institute of Statistics)
Managerial personnel feminization rate	GENMAN	ISTAT
Human resource turnover index	HUMREP	(Italian National Institute of Statistics)
Human resource turnover rate	HUMTUR	ISTAT
Managerial personnel to personnel ratio	INCMAN	(Italian National Institute of Statistics)
Personnel under 35 years of age to total personnel ratio	PERYOU	ISTAT
Personnel with less than 20 years of seniority to total personnel ratio	SENPER	(Italian National Institute of Statistics)
Diploma-holding or higher-educated personnel to total personnel ratio	TRDEMPAL	ISTAT
Graduate or higher-educated personnel to total personnel ratio	TRDEMPDE	(Italian National Institute of Statistics)
Graduate or higher-educated managerial personnel to total managerial personnel ratio	TRDMANDE	ISTAT
Average age of workers	AGE	State General Accounting Office
The total number of employees within each organization	TOTAL EMPLOYEES	State General Accounting Office
Percentage of graduate workers	EDUCATION	State General Accounting Office
Percentage of employees with more than 20 year of seniority	SENIORITY	State General Accounting Office
	AMOUNT	OPENCOESIONE
	IQI	

On equilibrium efficiency in non-life insurance market

Giuseppe De Marco, Christian Di Pietro, Mariafortuna Pietroluongo [1,2]

Abstract

The paper investigates the dynamics of equilibrium efficiency in the non-life insurance market, emphasizing the role of cooperation in shaping premium formation. The premium is influenced by two key factors: the external effect of competitors premium and policyholders' behavior: latter can react to a change in premium by choosing another insurance contract. Regulators can analyze insurers' behavior – whether it is competitive or collusive– by observing their adjustments to premiums in response to regulatory constraints. This behavior reveals the underlying strategies insurers adopt to balance profitability and market dynamics. In details, the study adopts the standard model, in which insurers choose the premium price in order to maximize the expected profit. The competition case is analysis by Taylor (1986) and (1987), Polborn (1998), and Dutang et al. (2013), here we propose a concept of collusive solution, which is a refinement of Pareto optimal allocation.

JEL classiFIcation: C61, C70, G22, G30.

Key words: Non–Life insurance, Solvency constraint, Competitive Equilibrium, Pareto Optimum, Game theory.

We wish to thank AMASES 2022 conference and AMASES 2023 conference participants for comments. Special gratitude we wish to express to for stimulating the questions addressed in this paper. Any remaining errors are our own.

[1] Giuseppe De Marco: Universitá degli Studi di Napoli Parthenope, Dipartimento di Studi Aziendali e Quantitativi, Via Generale Parisi 13, 80132 Napoli, Italy. e-mail: giuseppe.demarco@uniparthenope.it;

[2] Christian Di Pietro, Mariafortuna Pietroluongo: Universitá degli Studi di Napoli Parthenope Dipartimento di Studi Aziendali ed Economici, Via Generale Parisi 13, 80132 Napoli, Italy. e-mails: christian.dipietro@uniparthenope.it, mfortuna.pietroluongo@unparthenope.it

1 Introduction

The insurance market operates under inherent uncertainty, primarily due to the unpredictability of the total claim cost, which includes both the amount of the cost and the probability of the claim. This uncertainty drives the dynamics of competition and potential collusion among insurers: to mitigate the risk of bankruptcy, insurers may collude by setting premiums higher than the fair premium. This cooperative behavior ensures higher revenues to buffer against unexpected losses. On the flip side, insurers strive to attract new policyholders and retain existing ones by adjusting premiums competitively. Policyholders respond to change in premium of insurance by signing up a new contract with another insurance company. The insurer's expected profit is subject to the risk profile of policyholders in his portfolio; as in Rothschild and Stiglitz (1976) and Rees et al. (1999), we consider an asymmetric problem information for the insurers about the policyholder's risk profile, but insurers are aware of the main characteristics of the population risk distribution. The possibility of separating contracts, as in Rothschild and Stiglitz (1976), is binding, and as in Willignton and Alegrìa (2012), consumers are obliged to underwrite an insurance contract. In a model such as this, as Wambach (1999) observed, we could fall into the Bertrand paradox, in which two firms are sufficient to obtain a zero-profit condition, and the consequence is a premium equal to marginal cost. Willignton and Alegrìa (2012) analyze the primary market featured by capac- ity constraints; Sonnenholzen and Wambach (2014) prove that the Bertrand paradox might not hold true in the presence of firm capacity constraints and firms charge a price above marginal costs. Observe that capacity constraints could help insurance companies to achieve collusion. Collusion can take various forms, such as price fixing, market allocation, or production quotas. Trivially, under collusion behavior, the Bertrand paradox vanishes.
Another way to escape Bertrand's paradox, the one used in this paper, is using the elasticity of demand parameters, which represent the response of the underwriting strategy to a reduction or an increase of 1% by policyholders with respect to the change in premium. From the insurance market, the elasticity of demand can be interpreted as the level of trust of policyholders in the future coverage of claims by insurers, the elasticity of demand will be finite which means that the premium is not the only discriminating factor for the policy- holders underwriting strategy. According to this interpretation, the elasticity of demand should influence collusive behavior; we expect that the higher the elasticity, the lower the possibility of collusive behavior, in the sense that a careful regulator will sanction it.

We examined the insurance market through the lens of an n-person game, where insurers use the best underwriting technique to maximize predicted profit (also see Dutang et al. (2013), Taylor, (1986), (1987), and Polborn, (1998)). The insurers in Klinger and Levikson (1998) face the identical optimization problem, but without competitors' effects on the demand function. Insurers assume the policyholder's risk at this point and demands to settle claims are made again. On the other hand, it is common knowledge that insurance companies adopt cooperative strategies to improve market stability and prevent excessive fluctuations among policyholders. Indeed, it is well known that insurance companies cooperate with each other in managing risk through reinsurance contracts in the secondary market. Cooperative behavior is tolerated in the secondary market after companies have offered their contracts to customers; the reinsurance contracts must be Pareto-optimal; see, for example, Borch (1962) and Ghossoub et al. (2022). As write in the abstract of Borch (1962), "if each participant seeks to maximize his utility, taking this price as given, the market will not in general reach a Pareto optimal state. If the market shall reach a Pareto optimal state, there must be negotiations between the participants, and it seems that the problem can best be analyzed as an n-person cooperative game", reinsurance serves as a critical risk management tool for insurance companies, enabling them to protect themselves against significant financial losses and to expand they're under capacity. Although interesting, the optimal price bargaining of reinsurance contracts is not the subject of study. We are interested in whether collusive behavior, negotiations, may exist between insurance companies when they decide the premiums.

We study in a game theoretical framework the insurance market in which insurers can collude during the premium formation and show whether and how the regulator can gather information on the existence of collusion by looking at insurers' reaction. The collusive equilibrium is also a Pareto optimum for the insurance market, studies show the bargaining to transfer risk to achieve a Pareto Optimum. As observed by Willignton and Alegrìa (2012), a collusive seeming outcome may occur into the primary market featured by capacity constraints; bargaining between insurers helps the market to contrast the negative effect of default caused by capacity constraints. To make the study clear, we introduce the concept of Kantian equilibriums given by Roemer (2010). As we will see, the set of Kantian equilibria that arise in this context includes the set of Pareto optima. According to Roemer's (2010) definition of Kantian equilibrium, players in a strategic context choose their actions based on a principle of universalizability, rather than restrictively optimizing their individual payoffs. This means that each player evaluates their choice by considering the collective outcome that would result if everyone were to follow the same decision rule. In contrast to Nash equilibrium, where players take others', strategies as given and act in their own self-interest, a Kantian equilibrium ensures that no player will want to unilaterally deviate if it means violating a principle that would lead to a worse collective outcome when universally applied. A classic example of Kantian reasoning in economics is cooperative behavior in public goods provision: instead of free-riding, players might contribute because they consider the world where everyone contributes, leading to a better overall outcome.

The concept of Kantian equilibrium offers another way to understand how agents might overcome the tragedy of the commons. Instead of optimizing selfishly (as in Nash equilibrium), they ask: What if everyone followed the same principle when deciding their economic actions? If they recognize that overproduction leads to collective ruin, they may all choose to reduce effort in a way that sustains the economic's productivity. Roemer (2010) shows that under Cournot competition, each collusive equilibrium is a Kantian equilibrium and vice versa. Here we prove that under Bertrand competition this statement is no longer true.

The paper is a first step to understanding the regulatory challenges when insurance company form the premium. Our research specifically investigates on equilibrium outcomes, if regulatory constraints modifying an existing equilibrium or creating a new one. The introduction of constraints by regulators can create a new equilibrium between the parties involved, understanding how these changes occur and how they can affect the insurance market is crucial for predicting the consequences of actions and guiding decision-making. We wish to investigate in which way the elasticities parameter and the equilibrium, competitive or collusive, are connected. Boonen (2016) sheds light on the role the regulator can play when companies contract for risk redistribution.

The paper is organized as follows: in section 2 we present the game structure, section 3 the regulator's choice about constraints, section 4 equilibrium concepts and results about the existence and uniqueness of a premium equilibrium. Section 5 conclusion and perspectives are given.

2 The *I*–person game

The non–life insurance market is modeled as an *I*–person game, the *I* insurers compete to serve n consumers, with n is large relative to *I*. Each insurer aims to maximize their expected profit by employing an optimal underwriting strategy. When insurers compete aggressively, Bertrand competition emerges, where each insurer sets premiums to maximize profits while considering competitors' pricing strategies. For foundational work, see Taylor (1986, 1987). The premium vector is represented by the vector $\boldsymbol{x} := (x_1, \dots, x_I) \in \mathbb{R}^I$ where the *j*–coordinates is the premium chosen by the *j*–insurer. The set of feasible premiums for insurer j is indicated with D_j and is affected by the regulator's choice. In this model we take into account the externality of each competitor on the expected profit of insurer, as a consequence on the game structure we have that all profit functions are defined on $\Pi_{j\in I}\ D_j$.

Each policyholder can underwrite a single insurance contract, this ensuring no overlap in coverage across insurers and each consumer is obliged to underwrite an unique contract, so each consumers is a policyholder.

In line with Dutang et al. (2013), the expected profit for the j–insurer is the product between its demand function and its expected profit for a single contract, as follows:

$$O_j(x) = D_j(x)(x_j - \pi_j) = \frac{n_j}{n}\left(1 - \beta\_j\left(\frac{x_j}{m_j(x)} - 1\right)\right)(x_j - \pi_j)$$

Where $\pi_j := \omega_j a_{j0} + (1 - \omega_j)m_0$ is the break–even premium for j–insurer, a_{j0} is the actuarial premium based on the pass experience of j, m_0 is the market premium and $\omega_j \in [0, 1]$ is the credibility factor. The break-even is a weighted sum between two estimated costs, one given by past experience and the other by the market. The elasticity parameter associated with the j–insurer is indicated with β_j, and represents the response of the underwriting strategy to a x% reduction or increase from policyholders.
We assume that the demand function is neither perfectly elastic nor perfectly inelastic $0 < \beta_j < +\infty$. It is a measurement by policyholders of trust, or, better said, a measure of the level of trust for the claim's coverage by the insurance company. The regulator knows the elasticity parameter of each insurance company.
The external effect for the premium formation is given by the market premium proxy $m_j(x)$, which is equal to the average of the competitors premium

$$m\_j(x) = \Sigma_{i \neq j} x_i/(I - 1)$$

The initial portfolio size is $(n_1, \ldots, n_I) \in \mathbb{R}^I$, n_j is the number of contract underwritings from j–insurer in the last year; it is obvious that the total portfolio size is $\Sigma_{j \in I} n_j = N$.
Each insurer's expected profit function is strictly concave for strictly positive premium vector.

3 The regulator

The regulator's primary goals are to ensure fair and competitive prices for policyholders and to promote a healthy competitive environment. The regulator plans the set of constraints for the premium formation in order to mitigate market inefficiency caused by collusion or excessive competition that could lead to financial instability for insurers or unfair outcomes for consumers. Solvency II drives us to a constraint with respect to the aggregate portfolio risk. The regulator imposes a solvency function for each insurance company in the

market. The solvency function for j–insurer is given by the formula:

$$g_1(x_j) = \frac{K_j + n_j(x_j - \pi_j)(1 - e_j)}{k\sigma(Y)\sqrt{n_j}}$$

The terms that define the Solvency function are: the numerator corresponds to the sum of the current capital K_j and the expected profit on the in-force portfolio of j–insurer, $e_j \in (0, 1)$ is the expense rate for a single contract; and the denominator approximates the required capital, where k is the solvency coefficient chosen to approximate a 99·5% quantile.
From the mathematical point of view, the function defined above makes sense for $n_j > 0$; moreover, for each insurer, we can define a premium threshold

$$x_j^g = \frac{k\sigma\sqrt{n_j} - K_j}{n_j(1 - e_j)} - \pi_j$$

We have a threshold described by the Solvency II constraint for each insurer. The Solvency II constraint function defines a lower bound for each insurer; for the above discussion, insurers with small portfolio sizes or high capital don't suffer the solvency as a real bound for the premium formation.

In addition to Solvency II, the regulator sets two functions that define the minimum premium $\underline{x}$ and the maximum premium $\overline{x}$ for the market. In formula, we have

$$g_2(x_j) = x_j - \underline{x} \qquad g_3(x_j) = \overline{x} - x j$$

The premium of j–insurer is called feasible when belongs to the following compact set: $\Pi_{j\in I}[max\{x_j^g,\ \underline{x}\}, \overline{x}\ ]$.

4 Equilibrium concepts

Once the constraints and expected profit functions for each insurer are established, it becomes possible to define the key equilibrium concepts for the insurance market. These include competitive equilibrium, collusive equilibrium, and Kantian equilibrium, which provide insights into how insurers behave un- der varying competitive and cooperative dynamics.

Definition 1 *The vector* $x^* = (x_1^*, \ldots, x_I^*)$ *is a competitive–equilibrium for the insurance market if for all* $j \in I$ *the premium* x_j^**solves the following maximization problem:*

$$\max_{x_j} O_j\left(x_j, x_{-j}^*\right)$$

$$s.t.\, g_l\left(x_j\right) \geq 0 \;\; l = 1,2,3$$

Where x_{-j}^* *is the vector premium chosen by competitors of* j.

In Rotshild and Stiglitz the demand function has perfect elasticity, so the stable configuration satisfies the zero-profit condition; here, the optimal premium is not defined by the zero-profit condition.

Proposition 2 *If the parameter elasticity* β_j *is finite, then the premium vector* $(\pi_1, \ldots, \pi_I)$ *is not an equilibrium for the Insurance market.*

As a comment to the above result, the parameter elasticity β_j measures the trust that insurers have in the insurer's ability to cover future claims. From a policyholder's point of view, an insurer who sells the contract at a too low premium relative to the current contract may not have enough capital to cover the claims in the future. The insurer's reputation becomes important when selling contracts for the policyholder's optimal choice. As an effect of the insurance market, insurance contracts are not perfectly substitutable, so we conclude that the market is characterized by monopolistic competition.

Dutang et al (2013) prove the existence of an unique competitive–equilibrium for the constrained problem as given in Definition (1) here we sum up the result.

Proposition 3 *The constraint problem as given in definition 1 admits a unique equilibrium.*

Insurance companies adopt cooperative strategies to improve market stability and prevent excessive fluctuations among policyholders. Indeed, it is well known that insurance companies cooperate with each other in managing risk through reinsurance contracts. In this model the collusive equilibrium arises from a collective behavior during the premium formation. We set the collusive equilibrium as the argmax of the aggregate expected profit of insurers.

Definition 4 *The vector* $x^* = (x_1^*, \ldots, x_I^*)$ *is a collusive equilibrium for the* *insurance market if x^* solves the following maximization problem*

$$\max_{x_j} O^s\left(x_j, x_{-j}^*\right)$$
$$s.t.\, g_l\left(x_j\right) \geq 0 \;\; l = 1,2,3$$

we define $O^s(x)$ as the sum of all the firm expected profit functions, in formula

$$O^s(x) = \Sigma_{j\in I} O_j(x)$$

Trivially by Definition (4), we get that all the collusive equilibrium falls under the broader concept of Pareto Optimum for the game $(I, \left(O_j\right)_{j\in I})$. A premium vector x is Pareto Optimum if there doesn't exist an alternative premium vector y such that some insurers state better off and there is at least one insurer who states worse off.

Proposition 5 *The collusive equilibrium is a Pareto Optimum*

For the constrained collusive solution, we apply the well-known Weierstrass theorem since the object function is defined on a compact set $\Pi_{j\in I} D_j$. Recall that if O^s is a concave function, the Weierstrass theorem guarantees the uniqueness of the maximum.

Proposition 6 *The constraint problem as given in definition 4 admits at least maximum. Moreover, if for all $i \in I$ holds inequality $\pi_i < max\{x_i^g, \underline{x}\}$ then the maximum is unique.*

For the unconstrained problem, the direct calculation of collusive equilibrium is not so simple. Therefore, we introduce the definition of Kantian equilibrium presented in Roemer (2010), in which any insurer postulates that other agents act with the same kind of action, i.e., an equiproportional charge of premium. A vector x belongs to $\mathbb{R}_+^I$ is a Kantian equilibrium if the restriction of function O^s on the one-dimensional vector space defined by the same vector x reach the maximum exactly in x. More formal we get.

Definition 7 *The vector $x^* = (x_1^*, \ldots, x_I^*)$ is a Kantian–equilibrium for the insurance market if $\alpha = 1$ solves the following maximization problem:*

$$\max_{\alpha} O_j(\alpha x^*)$$
$$s.t.\, g_l\left(\alpha x_j^*\right) \geq 0 \;\; l = 1,2,3$$

We use the concept of Kantian equilibrium cause is simpler to verifies respect the collusive equilibrium. Technically, the Kantian equilibrium is the maximum respect with a linear direction (variation in one dimension) while the collusive equilibrium is a local maximum variation in more than one dimension. Trivially a collusive equilibrium is also a Kantian equilibrium, let $x^* \in \mathbb{R}_+^I$ be a collusive equilibrium then $O^s(x^*) \geq O^s(\alpha x^*)$ for each $\alpha \in \mathbb{R}_+$. Therefore, when Kantian equilibrium does not exist, we can say the same about the collusive equilibrium.
To this aim we prove that the expected profit $O_j(x)$ is linear with respect to the variable α for each j. For that reason, the function $O^s(x)$ is also linear respect with α and it does not allow a maximum in the un- constraint problem.
We prove that the demand function $D_j(x)$ is a scala–invariant function with respect to α.

Proposition 8 *Let $x \in \mathbb{R}_+^I$ a premium vector and $\alpha \in \mathbb{R}_+$ then $D_j(\alpha x) = D_j(x)$ for all $j \in I$*

Then the expected profit $O_j(\alpha x) = D_j(x)(\alpha x_j - \pi_j)$ is linear with respect to the variable α, and its derivative is equal to $D_j(x)x_j$ which is null if and only if $x_j = 0$ or $D_j(x) = 0$. The former condition is exclude by assumption of positive premium vector, then a premium vector is Kantian if $D_j(x^*) = 0$ for all $j \in I$.

Proposition 9 *For the unconstraint case there are not Kantian equilibrium.*

We can conclude that without constraints there are not collusive equilibrium, there exists at least an insurer who finds it worthwhile to raise the premium. As observed by Roemer (2010), if the object function O^s is monotone increasing (or decreasing) with respect to all its arguments x_j, then each Kantian equilibrium is a Pareto optimum. He has proved that for oligopolists games under Cournot competition all the collusive solutions are also Kantian equilibrium and vice versa. Unfortunately for our model the collusive expected profit $O^s(x)$ is not monotone, so we cannot use the Roemer's results, but we know that a collusive equilibrium is also a Kantian equilibrium in the sense of definition (7).

We introduce the concept of Kantian equilibrium under constraints in which the set of feasible premiums is bounded. It has a lower bound and an upper bound; to define the Kantian equilibriums we use the truncation function $Z(\alpha x^*) = (\alpha x_i^*)_{i=1\ldots,n}$ as follow:

$$\alpha x_j^* = \begin{cases} \underline{x} \ \ if \quad \alpha x_j^* \leq \underline{x} \\ \alpha x_j^* \quad if \quad \underline{x} \leq \alpha x_j^* \leq \overline{x} \\ \overline{x} \ \ if \quad \overline{x} \leq \alpha x_j^* \end{cases}$$

Now we are ready to apply the Weierstrass theorem.

Proposition 10 *Take x^k a Kantian equilibrium, i.e. a premium vector which satisfies the definition 7, there exists at least $i \in I$ such that $x^k = \overline{x}$.*

Corollary 11 *Take x^* a collusive equilibrium, then there exists at least $i \in I$ such that $x^* = \overline{x}$*

5 Conclusion

In this paper, we compared two different steady-state concepts in the problem of premium formation by insurance companies. On the one hand, the well-studied Nash-style competitive equilibrium, and on the other hand, the Pareto optimum. In the insurance market, they have been studied separately—the competitive equilibrium when the agent transfers the risk to the company, and the Pareto optimum when the companies redistribute the risk among themselves. We depart from the model proposed by Dutang et al. (6), introducing, in addition to competition, a collusion equilibrium concept. The cooperative behavior leads us to a Pareto optimum, which refers to a maximum for the aggregate profit, in line with Golubin (8). Trivially, for the unconstrained case, equilibrium sets are disjoint, making collusive behavior recognizable. However, under the actions of a regulator, collusive and competitive behavior could define the same equilibrium. The analysis made in this paper may shed light on this grey zone, showing how elasticities affect equilibrium outcomes. We can conclude that under regulatory control, by the imposition of constraints, and knowing the fundamentals, one can understand the type of competition between insurance companies.

Moreover, the role of Solvency II is crucial in a competitive market, where the equilibrium exists but may lead to unstable risk coverage, whereas it is almost useless in a collusive market, where insurers do not lower the price to sell policies.

For the competitive case:

- In a context of low elasticity, the regulator, through its actions, guarantees an almost fair premium.
- In the presence of high elasticity, its role must be to mitigate the risk of insolvency.

In the non-life insurance market, the regulator is strongly oriented towards solvency problems rather than counteracting collusive behavior through fair premiums. Moreover, a Pareto optimum configuration for insurers represents an unfair premium for policyholders.

The equilibrium concepts considered in this paper help us to assess the level of competition in the insurance market. The Nash equilibrium is a natural concept when there is no collusion between insurers, whereas the Pareto optimum arises when the parties involved come to an agreement.

For future research, the proposed game models here can be extended:

- To a dynamic model of rational behavior by insurers,
- And to insurance markets involving two companies.

Appendix: proofs

Proof. Proposition 2
An unconstraint equilibrium solves the system

$$\begin{cases} \frac{n_1}{n}\left(1-\beta_1\left(\frac{2x_1-\pi_1-m_1(x)}{m_1(x)}\right)\right) \\ \vdots \\ \frac{n_I}{n}\left(1-\beta_I\left(\frac{2x_I-\pi_I-m_I(x)}{m_I(x)}\right)\right) \end{cases}$$

The premium vector x is an equilibrium for the unconstraint game if and only if for any $j \in I$ we have $m_j(x) = \frac{\beta_j}{1+\beta_j}(2x_j - \pi_j)$. Replace π_j instead of x_j and obtain that

$$m_j(x) = \frac{\beta_j}{1+\beta_j}\pi_j$$

for all $j \in I$. Since the elasticity β_j is strictly positive we get

$$\pi_j > m_j(x) = \sum_{i \neq j} \frac{\pi_i}{I-1}$$

this inequality must be satisfied for each $j \in I$, by trivial algebra these inequalities cannot be satisfied simultaneously.

■

Proof. Proposition 3
The object function is (strictly) concave. Moreover, for all insurers, the solvency constraint defines linear functions. We are in the hypotheses of Theorem 1 of Rosen (1965), the existence is guaranteed. As observed by Dutang et al (2013) the object functions are diagonally strictly concave. This condition is sufficient for the uniqueness.

■

Proof. Proposition 5

Suppose that x is a collusive equilibrium and assume it is not a Pareto Optimum, there exists y such that $O_j(y) \geq O_j(x)$ with the strictly inequality for some j. As a consequence, we get the contradiction $O^s(y) > O^s(x)$.

■

Proof. Proposition 8

Of market proxy definition we have $m_j(\alpha x) = \alpha m_j(x)$ for all $j \in I$. Take the demand function, we get

$$D_j(\alpha x) = D_j(x)$$

■

Proof. Proposition 9

Consider the objective function $O_j(x)$ and calculate the derivative respect with α. The derivative is null if and only if $x_j = 0$ or $D_j(x) = 0$ for all $j \in I$. It is sufficient to prove that for each j the premium vector x satisfies the equality $D_j(x) = 0$. It means that $x_j = \frac{\beta_j}{1+\beta_j} m_j(x) > m_j(x)$ for all $j \in I$ and it is absurd by the definition of market proxy.

■

Proof. Proposition 10

Indicate with x^k the Kantian equilibrium and suppose that $x_i^k < \overline{x}$ for all $i \in I$, so we can find α such that

$$1 < \alpha \leq \frac{\overline{x}}{\max_i \{x_i^k\}}$$

By the strictly linearity of the objective functions respect with α we know that $O_j(x) < O_j(\alpha x)$. That inequalities drive us to obtain the absurd.

■

References

[1] Albrecher, H. and D. A. Dalit, (2017), On effects of asymmetric information on non-life insurance prices under competition, *International Journal of Data Analysis Techniques and Strategies*, 9, issue 4, p. 287-299, https://EconPapers.repec.org/RePEc:ids:injdan:v:9:y:2017:i:4:

[2] Asimit, V., and T. J. Boonen, (2018). Insurance with multiple insurers: A game-theoretic approach. *European Journal of Operational Research*, 267(2), 778-790.

[3] Asmussen S., B. J. Christensen and J. Thøgersen, (2019). Nash equilibrium premium strategies for push–pull competition in a frictional non-life insurance market *Insurance: Mathematics and Economics*, vol. 87, 92–100.

[4] Boonen, T. J. (2016). Nash equilibria of over-the-counter bargaining for insurance risk redistributions: The role of a regulator. *European Journal of Operational Research*, 250(3), 955-965.

[5] Borch K., (1962). Equilibrium in a Reinsurance Market. *Econometrica* 3, 424–444.

[6] Dutang C., H. Albrecher and S. Loisel, (2013). Competition among non–life insurers under solvency constraints: A game–theoretic approach. *European Journal of Operational Research* 231, 702–711.

[7] Ghossoub M., W. Jiang and J. Ren, (2022). Pareto-optimal reinsurance under individual risk constraints. *Insurance: Mathematics and Economics* 107, 307-325, ISSN 0167-6687.

[8] Golubin, A. (2005). Pareto-optimal Contracts in an Insurance Market. *ASTIN Bulletin: The Journal of the IAA* (2), 363-378. doi:10.2143/AST.35.2.2003458

[9] Kliger, D., and B. Levikson, (1998). Pricing insurance contracts-an economic viewpoint. *Insurance: Mathematics and Economics*, 22(3), 243-249.

[10] Polborn M.K., (1998). A Model of an Oligopoly in an Insurance Market. *The Gineva Papers on Risk and Insurance Theory* 23, 41–48.

[11] Rees, R., H. Gravelle and A. Wambach, (1999). Regulation of Insurance Markets. *The Geneva Risk and Insurance Review*, 24(1), 55-68, https://EconPapers.repec.org/RePEc:pal:genrir:v:24:y:1999:i:1:p:

[12] Roemer, J.E., (2010). Kantian Equilibrium. *The Scandinavian Journal of Economics*, 112(1), 1-24.

[13] Rothschild M. and J. Stiglitz (1976) Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4) (Nov., 1976), 629-649.

[14] Sonnenholzner M. and A. Wambach, (2006). Oligopoly in Insurance

Markets. Chapter in **Encyclopedia of Actuarial Science**. DOI: 10.1002/9780470012505.tao003.

[15] Taylor G.C., (1986). Underwriting Strategy in a Competitive Insurance Environment. *Mathematics and Economics* 5, 59–77.

[16] Taylor G.C., (1987). Expanses and Underwriting Strategy in Competition. *Mathematics and Economics* 6, 275–287.

[17] Wambach A., (1999). Bertrand competition under cost Uncertainty. *International Journal of Industrial Organization* 17, 941–951.

[18] Warren, R., J. Yao, T. Rourke, and J. Iwanik, (2012). Game Theory in General Insurance: How to outdo your adversaries while they are trying to outdo you. *In GIRO Conf* (cit. on p. 69).

[19] Willington M. and A. Alegría (2012). Collusion in a One-Period Insurance Market with Adverse Selection, *The B.E. Journal of Economic Analysis & Policy*, 12(1).

[20] Wu, R., and A. A. Pantelous, (2017). Potential games with aggregation in non-cooperative general insurance markets. *ASTIN Bulletin: The Journal of the IAA*, 47(1), 269-302.

Imperfect Substitutability Between Old And Young Workers.

Salvatore Carrozzo* Alessandra Di Pietro†

Abstract

Employment rate of older workers in Italy has increased over the last decade, meanwhile youth employment rate have experienced a big decline. These divergent employment paths raise a question about the substitutability between old and young workers. In order to answer that question, we propose a novel identification strategy to estimate the elasticity of substitution in production between old and young workers. We start setting the labor demand functions for both groups within the same region-occupation-time group to estimate such elasticity. Then, we develop a theoretical model that shows towards-zero estimation bias induced by time correlations within each region-occupation-time group. To overcome this estimation problem, we use a set of instruments based on yearly employment changes by age and citizenship. Using yearly Italian administrative data for the period 1995-2004, we exploit a number of pension and labor migration reforms to create a set of exogenous instruments to time correlations within a region-occupation-time group. Finally, we find that old and young employees within the same region-occupation-time cell experience imperfect substitutability in production.

Keywords: Old-Young Elasticity; Long Run Adjustments; Labor Mobility

JEL codes: J21, J31, K31

*Corresponding author. Department of Business and Economics, University of Naples "Parthenope". Email: salvatore.carrozzo@collaboratore.uniparthenope.it.

†Department of Economics and Statistics, University of Turin and Collegio Carlo Alberto. Email: alessandra.dipietro@carloalberto.org.

1 Introduction

May increases in both life expectancy and retirement age affect youth employment opportunities? Most developed countries experience a decline in either youth employment rate or youth participation rate together with an increase in older participation rate. In 2018, France's and Italy's youth unemployment rates were still larger than pre 2008-crises, 20.08% and 32.2% (OECD database) respectively. While, the U.K.'s and the U.S.'s youth participation rates were 5% and 4% (OECD database) lower than pre 2008-crises, respectively. At the same time, all developed countries experience an increase in participation and employment rate for workers age over 55. These divergent patterns raise a question on the existence of a large degree of substitutability between old and young workers.

The substitutability between old and young workers is an outstanding question in labor literature. On the one hand, the *lump of labor* concept claims that old and young workers compete for a scarce good: a job. Boeri et al. (2017) show that the 2011 sudden increase in the retirement age of *baby-boomers*, the generation born between the end of the World War II and the late '50, due to a pension reform has negatively affected youth employment in Italy. Mohnen (2019), using 1980-2017 U.S. data, finds that the effect of an increase in the retirement age on the youth employment is wider the larger older worker share in low skilled jobs. Bertoni and Brunello (2017) find the same results in Italy between 2004 and 2015. Further, Bovini and Paradisi (2018) find that the effect of an increase in Italian retirement age on youth labor outcomes is wider the larger share of manufacturing workers over the period 2009-2015. On the other hand, the existence of imperfect substitutability between old and young workers should lower the competition for the same job. Brugiavini and Peracchi (2010) find that delaying retirement has a positive on the youth employment rate in Italy between 1997 and 2004. Gruber and Wise (2010) find a positive effect of an increase in older participation rate on youth employment rate by studying labor markets of several developed economies from late '70 to the beginning of the new century. Munnell and Wu (2012), using 1977-2011 U.S. data, show that an older employment increase leads to better labor outcomes for young workers, raising both wages and employment rate. Our paper fills in by offering a novel identification strategy to estimate the old-young elasticity of substitution in production.

In our paper, the elasticity of substitution between old and young workers is the ratio of the percentage change in old-young employment ratio (labor gap) to the percentage change in the old-young wage ratio (wage gap) within the same region-occupation-year cell. We estimate the inverse of such elasticity to identify the causal relation of an increase in labor gap on wage gap. The greater is the identified effect, the smaller is the substitutability between old and young workers. To put it in another way, imperfect substitutability leads to a smaller effect on the age-group wage not affected by the employment increase.

We estimate a structural model to find the elasticity of substitution between old and young workers. We choice a nested constant elasticity of substitution (CES) production function to derive the relation between labor gap and wage gap for two reasons. First, the nested CES dimensions, in our case region-occupation-age-year, allow to control for different demand shifts. Second, the linearity of log first order conditions enables to study the old-young elasticity of substitution by using linear estimators. To estimate the model, we use 1995-2004 Work Histories Italian Panel (WHIP) employee data to estimate such elasticity. We restrict the sample to 712,514 full-time male workers in the private sector as they experience larger employment spells and, hence, accumulate on-the-job human capital at a constant pace. We aggregate data on employees to build total employment and average wage per region-occupation-age-year cell.

Estimating the effect of a labor gap change on the wage gap is not trivial as long-run dynamics might bias the estimates. In the short run, age-specific labor supply shocks might lower the wage gap through a change in the labor gap, but the occurrence of general equilibrium adjustments restores the previous wage gap equilibrium in the long run. Hence, labor supply shocks might have a negative effect on wages in the short run, but a positive one through general equilibrium adjustments in the long run. The net effect might be null, showing inverse-elasticity estimates biased towards zero, since the two effects offset each other. We call the general equilibrium adjustment mechanism *offsetting mechanism*, because it offsets any wage disequilibrium in the long run. Since the *offsetting mechanism* is unobservable and positively correlated with the labor gap and wage gap, elasticity estimates are upward biased.

Our paper contributes to solve this puzzle with two main innovations. First, we develop a theoretical model to understand how the *offsetting mechanism* biases the estimates. The idea is that the *offsetting mechanism* begins to adjust the disequilibrium a year after the labor supply shock applies. Hence, we model *offsetting mechanism* as a function of past labor supply shocks. In order to reduce the bias of the past shocks, a good instrument is a shock at current year. To the best of our knowledge, this way to study the *offsetting mechanism* bias is a novelty in the elasticity of substitution estimation literature.

Second, we provide a novel set of instruments to estimate the elasticity of substitution. As mentioned above, we have to find an instrument uncorrelated with past labor supply shocks to identify such elasticity. One candidate is the labor gap in first differences, because first differences sweep away past trends. However, labor gap in first differences might be correlated with region-occupation-year unobservable heterogeneity. In order to avoid this endogeneity issue, we combine the age dimension, old and young, with the citizenship dimension, native and foreigners, to create four instruments. Each instrument is the ratio of yearly region-occupation-age-citizenship employment change to the total yearly region-occupation-age employment. We exploit a deeper dimension aiming at lowering the region-occupation-year bias. To strengthen our identification strategy, we exploit a set of age-citizenship specific reforms enacted in Italy between 1995-2004. Reforms were enacted to save the Italian social security system from default by increasing both the retirement age and the workers per retiree. The timing of the reforms is suitable to identify the elasticity of substitution since not-serial correlated employment changes lowers the bias with past region-occupation-year labor gap changes.

We find that an increase in the labor gap lowers the wage gap by around 16% within the same region-occupation-year labor market. Further, the effect corresponds to an elasticity of substitution around 6, because in our theoretical model the effect reflects the negative inverse of such elasticity. The elasticity of substitution value range starts from 0, perfect complementarity, to infinity, perfect substitutability, our findings are closer to zero than infinity showing an imperfect degree of substitutability between old and young workers. Our findings are in line with the existing literature on old-young elasticity of substitution (Borjas, 2003; Card and Lemieux, 2001; D'Amuri et al., 2010; Manarcorda et al., 2012; and Ottaviano and Peri, 2012) as scholars find very similar results for different countries in different periods.

We provide a set of robustness checks and sensitivity analyses to test our results. First, we change the weights used to estimate the elasticity, because different weights may lead to different point estimates (Borjas, et al. (2012)). We use wage gap variance as a weight in our baseline estimates, while we weight for total employment in every region-occupation-year cell to test our results. Results do not change. Second, we test our identification with the foreign employment instrument. This instrument is very common in the literature and it is widely used to estimate the old-young elasticity of substitution. We show that the instrument is weak in our specification. Third, we assume that labor supply shocks identify the effect, through our instruments, excluding region-occupation-year demand shocks. We use temporary laid-off workers as a labor demand shock instrument to check our assumption. The instrument is weak and estimates are not significant. Fourth, we evaluate whether our parameters of interest are time-varying. We interact with our instruments with a linear trend increasing the set of instruments from four to eight. The results do not change. Hence, our baseline specification is robust to time dimension.

Related literature.— The literature on old-young substitutability in production exploits demographic changes to understand the degree of complementary among several age groups. Freeman (1979) is one of the first to study the degree of substitution among workers belonging to different cohorts. He estimates the elasticity of substitution between *baby-boomers* and previous cohorts in the US. He finds that an increase in the young labor supply has a larger effect on younger workers' wage than on older workers' one. This result shows workers belonging to different cohorts are imperfect substitutes in production. Katz and Murphy (1992) extend the analysis by exploiting the industry level variability. They identify demand shocks with technological shifts and labor shocks with demographic cohort characteristics. They show that both have a role in setting out the degree of imperfect substitutability across different age groups. A further extension of Katz and Murphy (1992) is Card and Lemieux (2001), where they improve the accuracy of elasticity of substitution estimates by taking into account both time effects and cohort effects. The underlying intuition relies on different salary paths among cohorts over time. By exploiting "baby-boomer" shock in the U.S., Canada, and the UK in a nested constant elasticity of substitution, they are able to make cross-country comparisons of the results. They estimate an elasticity of substitution among different age groups in the range of 4 to 6 by proving that additional fixed effects play an important role to exclude any possible bias due to supply or demand shifts. Borjas (2003), D'Amuri et al. (2010), Manarcorda et al. (2012) and Ottaviano and Peri (2012) extend Card and Lemieux (2001) exploit foreign labor force as instrument to estimate old-young elasticity of substitution, but thier estimates do not show any significant difference.

All the mentioned scholars do not pay much attention to long-run effects, while Llull (2018) points out that demand for different labor inputs depend on past labor supply shocks[1]. He shows that human capital accumulation is one of the main drivers to adjust wage disequilibrium in the long run. Also Jaeger et al. (2018) show that firms anticipate the labor supply shifts by adjusting the capital level. These mechanisms happen when labor force increases are stable across years. However, they only focus on foreign labor supply shocks, while we address the long-run bias issue by taking into account native labor supply shocks as well. We provide an estimation strategy that complements the literature on old-young elasticity of substitution estimate and adds an other piece to general equilibrium adjustment puzzle. Our estimation method relies on Arellano and Bover (1995) who exploit the first differences to identify the long run parameter in a dynamic panel framework. The underlying intuition is the same, but we apply that in a static framework.

The article proceeds as follows. Section 2 describes the institutional background over the considered time span. Section 3 presents the theoretical framework. Section 4 shows the data and the descriptive statistics. Section 5 discusses the empirical strategy. Section 6 shows the results. Section 7 presents robustness checks. Section 8 concludes.

[1]People reshape their human capital accumulation after experienced labor supply shocks.

2 Institutional Background

At the turn of the 20th century, Italy has experienced a number of labor market reforms mostly tackling the supply side. Among others there were pension reforms and migration flows regulations. Depending on the type of reform, different cohorts of workers were involved.' What follows is a brief review of all these different policies, grouped by theme.

2.1 Pension Reforms

The age threshold defining the active population of a country clearly affects the size of labor force, and pension reforms play an important role in setting such a threshold. The idea behind reforms in the '90s was keeping older workers in the labor market as longer as possible. This is due to an increase in life expectancy and experts were casting doubts on the sustainability of a pay as you go pension system. Two main pension reforms characterize the end of the century, Dini reform in 1995 and Prodi reform in 1997. As mentioned, the main aims were containment of public spending and curbing early retirement. The very first attempt to postpone retirement age (gradually) occurs with Amato reform, in 1992. In 1995, Dini reform, (L.335/1995), raised the age and contribution requirements for seniority pension. The change was gradual and finished in 2008. Prodi reform in 1997 further increases age and contribution requirements for seniority pensions.

2.2 Migration Reforms

Italy has long been a country of emigration. First regulations on immigration flows date back to the 80s. Up to that moment legalization of immigrant workers mainly happened through amnesties. In the '90s, a pool of migration laws enacted and included an amnesty to legalize migrant workers who had been working (or living) in the country for a year before. In 1990, Martelli law, L. 39/1990, was the first to regulate economic immigration in the country and legalize 215,000 foreign workers. This law imposed restrictions to incoming flows, and set a maximum number of workers to be accepted each year, based on foreseen Italian labor market needs. Dini decree, in 1995, allowed for the legalization of 244,500 immigrant workers. In 1998, the Turco-Napolitano law (L 40/1998) implemented major changes. This law represented the milestone for migration regulation in Italy. It involved inclusion of migrant workers within the labor force and made procedures and rules smoother and clearer. It allowed immigrants a temporary visa through the sponsorship channel to look for a job. Together with this reform, other 217,000 workers were regularized. Political debates spurred by increased migration flows conveyed into the Bossi Fini law (L. 189/2002). That stopped the sponsorship system and introduced stricter limitations to immigration. Few months after the enactment, the situation in black market was dramatic and, therefore, the two Ministers,Bossi and Fini, promoted the largest amnesty in Europe (634,700 immigrant workers were legalized). As such, some scholars used it to better understand the impact of amnesties on labor market outcomes. Devillanova et al. (2014) exploit it as a natural experiment and show that an increase in employment probability follows the prospect of legal status. Size of this increase is two third of the increase in employment rate illegal immigrant experience in the five years after entering the country. Di Porto et al. (2019) show the short term impact of 2002's regularization, with most of the legalized workers staying in the legal labor market for long. Amnesty regularized 62% of regular immigrants in the country in 2002 (Barbagli et al., 2004).

3 Model

3.1 Theoretical Framework

In order to study the elasticity of substitution between old and young workers we use a nested constant elasticity of substitution (CES) approach. Most of the prominent studies (e.g., Card and Lemieux, 2001; Borjas, 2003; Ottaviano and Peri, 2012) have used an aggregate production function to estimate the elasticity of substitution between old and young workers. An aggregate model provides an overview of national labor market but loses information about differences among local labor markets. We prefer to add the regional dimension to take into account local differences in labor force. We assume an identical Cobb-Douglas production function in each region r at time t:

$$Y_{rt} = A_{rt} K_{rt}^{\alpha} L_{rt}^{1-\alpha} \tag{1}$$

where Y is the output, A is exogenous total factor productivity, K is the physical capital, L is a CES aggregate of different types of labor, and α is the income share of capital. L_{rt} includes workers who differ by occupation and age, respectively. Let

$$L_{rt} = [\theta_{rBCt} L_{rBCt}^{\frac{\sigma-1}{\sigma}} + \theta_{rWCt} L_{rWCt}^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where BC (WC) indicates blue collar workers (white collar workers) and σ is the elasticity of substitution between blue collar and white collar workers ($0 \leq \sigma < \infty$). The θs are the region-occupation-time specific productivity parameters, with $\theta_{rBCt}+\theta_{rWCt}$=1. Finally, every occupation-specific labor input is a CES aggregate of imperfect substitute age-specific labor inputs. In particular,

$$L_{rst} = [\gamma_{rsOt} L_{rsOt}^{\frac{\lambda-1}{\lambda}} + \gamma_{rsYt} L_{rsYt}^{\frac{\lambda-1}{\lambda}}]^{\frac{\lambda}{\lambda-1}} \quad s = BC, WC \tag{3}$$

where O (Y) indicates old worker (young worker) labor input and λ, our parameter of interest, is the elasticity of substitution between old and young workers, with $\lambda \geq 0$. γs are the region-occupation-age-time specific productivity parameters, with $\gamma_{rsOt}+\gamma_{rsYt}=1$. We get the old (young) labor demand within each region-occupation-year cell by assuming that marginal product of old (young) labor is equal to the old worker (young worker) wage. Using logs, the age specific labor demand within each region-occupation-year cell is equal to:

$$ln(w_{rsat}) = ln(A_{rt} K_{rt}^{\alpha} L_{rt}^{-\alpha}(1-\alpha)) + \frac{1}{\sigma} ln(L_{rt}) + ln(\theta_{rst}) - (\frac{1}{\sigma} - \frac{1}{\lambda}) ln(L_{rst}) + ln(\theta_{rsat}) - \frac{1}{\lambda} ln(L_{rsat}) \tag{4}$$

Taking the difference side by side of labor demands for old and young workers, we get rid of all terms on the right-hand side, but the difference between productivity parameters and the difference between old and young labor demands:

$$ln(\frac{w_{rsOt}}{w_{rsYt}}) = ln(\frac{\theta_{rsOt}}{\theta_{rsYt}}) - \frac{1}{\lambda} ln(\frac{L_{rsOt}}{L_{rsYt}}) \tag{5}$$

Hereafter, we define the wage difference between the old and young worker as "wage gap" and the employment difference between old and young workers as "labor gap".

3.2 Labor gap between old and young by citizenship

In this subsection, we show how local and foreign labor supply shocks affect the labor gap between old and young workers. We assume that the employment level for old and young workers within the same region-occupation-year cell is a function of native and foreign employment[2] :

$$L_{rsat} = f(L_{rsaNt}, L_{rsaFt})) \tag{6}$$

where N (F) indicates native (foreigner) characteristics. We assume that each age-specific labor input is continuously differentiable and the first derivative with respect to citizenship dimension is greater than zero. Looking at differential[3] in discrete form we

[2]A wide range of literature (e.g. D'Amuri et al., 2010, Manarcorda et al., 2012 and Ottaviano and Peri, 2012) use a CES aggregate of native and foreign labor inputs in every age-specific group. We do not assume any functional form to avoid any constraint on the age-citizenship elasticity of substitution.

[3]In Appendix for the mathematical derivation.

obtain:

$$\Delta(ln(L_{rsOt}) - ln(L_{rsYt})) = \Sigma_c(\beta_{Oc}\frac{\Delta L_{rsOct}}{L_{rsOt}} - \beta_{Yc}\frac{\Delta L_{rsYct}}{L_{rsYt}}) \qquad c = N, F \tag{7}$$

where

$$\beta_{Oc} = \frac{\partial L_{rsOt}}{\partial L_{rsOct}} \qquad \beta_{Yc} = \frac{\partial L_{rsYt}}{\partial L_{rsYct}} \tag{8}$$

with c indicating if they are either natives or foreigners[4]. Each β is assumed to be fixed[5]and measures the elasticity of labor gap differential change with respect to a specific subgroup differential change, where a subgroup is the total number of workers with age a and citizenship c.

This decomposition allows us to understand how labor supply shocks affect the labor gap. By imposing β_{OF}=β_{YF}=β_F and rearranging the terms we get:

$$\frac{\Delta(ln(L_{rsOt}) - ln(L_{rsYt}))}{\frac{\Delta L_{rsOFt}}{L_{rsOFt}} - \frac{\Delta L_{rsYFt}}{L_{rsYFt}}} = \beta_F + \beta_{ON}\frac{\frac{\Delta L_{rsONt}}{L_{rsOt}}}{\frac{\Delta L_{rsOFt}}{L_{rsOFt}} - \frac{\Delta L_{rsYFt}}{L_{rsYFt}}} - \beta_{YN}\frac{\frac{\Delta L_{rsYNt}}{L_{rsYt}}}{\frac{\Delta L_{rsOFt}}{L_{rsOFt}} - \frac{\Delta L_{rsYFt}}{L_{rsYFt}}} \tag{9}$$

The left-hand side is the elasticity between a change in old-young labor gap and a change in old-young foreigner labor gap, where the denominator represents a change in foreign employment. We show that the effect of a foreign labor supply shock is not constant, but it depends on old and young native labor supply change.

This is an important result, in the literature a very common assumption is the exogeneity of foreign labor supply shock with respect to native employment. In our model, instead, the effect of a foreign labor supply shock on old-young labor gap depends also on native labor supply shocks (if any).

In the Empirical Strategy Section we take into account this finding to estimate the elasticity of substitution.

4 Data

The empirical analysis is based on the information on the wages and employment of old and young workers drawn from the 1995-2004 Work Histories Italian Panel (WHIP) dataset. The WHIP also contains information on citizenship that we use to create our set of instruments.

The WHIP database includes information on social securities records of 2,164,829

[4]This methodology is very similar to one developed by Amiti et al. (2019) that provide a decomposition of a firm price differential.

[5]In the empirical section we allow parameters to vary over time.

employees from 1995 to 2004, around 140,000 observations per year. Since our aim is to provide with a new identification strategy using a standard theoretical framework, we follow the literature to create wage and employment variables.

The main sample is restricted to men aged 18-64 working in the private sector. We narrow the analysis only to male employees as old and young females do not accumulate constantly experience in their working life showing larger substitutability (Freeman, 1979). Further, foreign females in Italy are usually employed as either caregivers or domestic helper, while native females are often employed in the public sector (Venturini and Villosio, 2008). We narrow the analysis only to private sector as public sector has more rigid labor dynamics. EU15 foreign workers are excluded from the analysis to avoid confounding effects due to similarities between EU15-foreign and Italian workers. Including EU15 workers in native (foreign) group might overestimate (underestimate) the elasticity of substitution. Further, the EU15 worker exclusion allows us to focus our attention on foreigners with different skills with respect to natives. We take the gross average log daily wage as wage measure[6]. Unfortunately, we have only information on total contribution days, where one contribution day is equal to 8 hours spent at work. The lack of information on hours spent at work does not allow us to include part time jobs as they differ by hours spent at work per day. Due to this missing information we are not able to homogenize the daily wage of part-time workers. Hence, we prefer to consider only full time workers. To measure the cell-specific employment we, first, measure the days spent in each region-occupation-age-year cell to the total worked days in a year per every worker. Then, we sum these shares to obtain the total employment in each region-occupation-age-year cell. We follow this procedure to overcome the assignment of a single worker to different region-occupation-age-year cells since there are some workers that change either working region or job within a year. Following the literature (i.e. Card, 1999), we assign the 'old' label for workers age over 38[7].

Table 1 and Table 2 show the shares in each macro-region-age-citizenship[8] cell by occupation. Table shows that for every 10 young blue (white) collar workers, there are 5 old blue collar workers (7 white collar workers). This evidence shows that old employees are more likely to hold a white collar occupation than a blue collar one. Further, the age-specific ratio of the blue collar total employment to the white collar total employment is

[6] We get rid of the first and last percentile of the distribution in order to avoid any confounding effect due to outliers.

[7] The explanation is that workers accumulate on-the-job human capital with a lower pace when they are age over 38.

[8] I show macro regions instead of regions for the sake of table clarity. The regional summary statistics lead to the same descriptive evidences. North includes: Emilia Romagna, Friuli Venezia Giulia ,Lombardia, Liguria, Piemonte and Valle d'Aosta, Trentino Alto Adige, Veneto. Centre includes: Lazio, Marche, Toscana and Umbria. South and Islands include: Abruzzo, Basilicata, Calabria, Campania, Molise, Puglia, Sardegna and Sicilia

equal to 2.96 and 2.08 for the young and old employees, respectively. These ratios show that old employees compete much more for white collar occupations since there are three young blue collar workers for each young white collar worker, and only two old blue collar workers for each old white collar worker. Hence, an increase in retirement age has a larger effect on young white collar workers than young blue collar workers. Further, Table 1 and Table 2 show that foreigners are more likely to be young and blue collar workers. Hence, an increase in foreign workers affects mostly the native young blue-collar workers. Table 3 and 4 show log daily wages for blue and white collars, respectively. In each table, we have information on log daily wages in each macro region-age cell over 1995-2004 period. Wages are deflated to 2001 euro by using OECD's Italian CPI series. We observe an overall sharp increase of real wages until 1999, followed by a decline and a renewal until 2004. Looking at the difference between the old and young average log daily wages, we see a declining path for blue collar wage differences, while there is no evidence of such trend for white collars. This finding suggests that old and young white collar workers are more substitute than old and young blue collar workers. Young blue collar wage does not respond to migration and pension reforms, while old blue collar wages lower over time by decreasing the gap with young blue collar workers. Instead, young white collar workers do not fill the wage gap with old white collar workers over time, showing a larger substitutability. Although these evidences are only descriptive, we might consider them as a first evidence on the degree of substitutability between old and young workers.

Finally, Figure 1 shows the trend for each instrument from 1985-2004. Each instrument is the ratio of yearly region-occupation-age-citizenship employment change to the total yearly region-occupation-age employment. We have four subgroups: old natives, young natives, old foreigners and young foreigners. Yearly changes are very noisy from 1995 to 2004, they do not follow a common path as observed for young and old natives in the previous periods. Hence, we exploit this variability to estimate old-young elasticity of substitution.

5 Empirical strategy

In this section we explain how to implement our theoretical findings in an empirical setting to estimate the elasticity of substitution between old and young workers. We estimate the equation (5) in Section 3. As shown in Section 4, we use as dependent variable the difference between average log daily wage of old workers and average log daily wage of young workers for each region-occupation-year cell (wage gap). Independent variable is the difference between the total employment of old and young workers for each region-occupation-year cell (labor gap). The old-young elasticity of substitution, $1/\lambda$, may be biased towards zero since labor gap might be positively correlated with long run *offsetting mechanism*. In the next subsection, we explain how long run adjustment may bias our estimates.

5.1 The *offsetting mechanism* function

We start estimating the equation (5) by substituting the log of old-young productivity ratio with a broad set of fixed effects and an error term. The new estimating equation is:

$$ln(\frac{w_{rsOt}}{w_{rsYt}})) = \phi_{rt} + \phi_{rs} + \phi_{st} - \frac{1}{\lambda} ln(\frac{L_{rsOt}}{L_{rsYt}}) + \varepsilon_{rst} \tag{10}$$

where ϕ_{rt} and ϕ_{st} capture every time productivity shift in each regional and occupation labor market, respectively. ϕ_{rs} captures any time invariant characteristic in each region-occupation labor market. ε_{rst} stands for whatever time variant residual components in every specific region-occupation labor market. However, we are not able to fully control for productivity shifts by using pairwise region-occupation-year fixed effects.

The unobservable region-occupation-year productivity shifts nested in the error term are very likely to be correlated with the labor gap. Indeed, the current labor gap and the current productivity shifts are both an increasing function of past age-group specific labor supply shocks. However, they have different effects on wage gap. An increase in the labor gap should have a negative effect on the wage gap, since an age-group specific labor supply has a larger negative effect on their own wages than on other age-group wages. On the other hand, an increase in productivity should have a positive effect on wage gap, since an increase in productivity might increase the wages. Hence, the positive correlation between productivity and both labor and wage gap might biases the estimates of the elasticity of substitution towards zero. We call this mechanism *offsetting mechanism* because it offsets the effect of the labor gap increase on the wage gap triggered by a labor supply shocks.

We want to shed light on this mechanism and on the correlation between labor gap and the *offsetting mechanism*. We assume that labor gap has a stable AR(1) process, we can see this process as MA(∞) process:

$$ln(\frac{L_{rsOt}}{L_{rsYt}}) = \sum_{k=1}^{t} \beta^t \epsilon_{rsk} + ln(\frac{L_{rsO0}}{L_{rsY0}}) \tag{11}$$

where $\epsilon_{rsk} = \Delta ln(\frac{L_{rsOk}}{L_{rsYk}})$, the last term is the yearly variation in the labor gap. Hence, on the right-hand side of Eq. (11) we have the sum of all labor gap yearly variations until t plus the initial condition. When $\Delta ln(\frac{L_{rsOk}}{L_{rsYk}}) \neq 0$, there is a change in the wage gap and the *offsetting mechanism* starts affecting later in time. Assuming that *offsetting mechanism* fully offsets wage shift in the following period[9], we nest the *offsetting mechanism* process in the error term:

$$\varepsilon_{rst} = \xi_{rst} + f(\epsilon_{rst-1}) \tag{12}$$

where ξrst is a random effect uncorrelated with the labor gap and $f(\epsilon_{rst-1})$ is the *offsetting mechanism* that depends on lag of yearly labor gap variation, ϵ_{rst-1}. As a result, this *offsetting mechanism* has a positive correlation with the labor gap biasing old-young elasticity of substitution estimate[10].

Offsetting mechanism function takes into account only the previous period labor gap change and not the current one. This structure allows us to exploit the current change as exogenous to the *offsetting mechanism* function.

The described mechanism is in line with papers that use new wave of migrants as an instrument to estimate the old-young elasticity of substitution (e.g. Borjas, 2003; Ottaviano and Peri, 2012), since yearly labor supply shocks are uncorrelated with previous ones[11].

5.2 A short run instrument

In the previous subsection, we discussed the structure of the *offsetting mechanism* process nested in the error term and the features that an instrument must have in order to identify the elasticity parameter. In our setting, we exploit variations based on current foreign and native labor supply shocks. We exploit the elements on the right-hand side of Eq. (7) as instruments, where each of them takes into account yearly change in each region-occupation-age-citizenship cell[12]. This methodology is very common in the macro literature, especially in dynamic panel data models[13].

In order to identify the true parameter and rule out all the time correlations between the instruments and the error term, we assume that employment time series in every subgroup is a random walk:

$$\Delta L_{rsact} = \varepsilon_{rsact} \tag{13}$$

[9]Adding more lags results hold.

[10]Because the $Cov(ln(\frac{L_{rsOt}}{L_{rsYt}}), f(\epsilon_{rst-1})) > 0$

[11]Jaeger et al. (2018) point out that the exogeneity of migration inflows depends on previous wave of migrants. If the flow of new migrants is stable across years, the labor demand can foresee the new inflow and adjust itself before it comes up.

[12]In the Appendix B, we compute parameter distortion when we use the first order differences as instrument and we assume that the *offsetting mechanism* function is linear. The estimated distortion is very small and equals to $-\frac{T}{(T-2)(T-1)}$, in particular it is smaller than the Nickell's one (Nickell, 1981).

[13]Arellano and Bover (1995) were the pioneers of this identification strategy that exploits the short run changes (i.e. first differences) to instrument levels. They use the lagged first differences as an instrument for the lagged value of the dependent variable that is their explanatory variable.

where, the first differences are equal to the error at current period. Testing this assumption we cannot reject the presence of unit root for every region-occupation-age-citizenship group[14].

The main concern is that the instruments might still depend on unobservable characteristics within region-occupation-year cell. To overcome this problem, we exploit Italian reforms enacted over the period 1995-2004. Between 1995 and 2014, Italy has enacted a series of national policies that have affected the labor supply of all considered categories across years. They provide us with exogenous variation to identify the effect as they are not labor market specific. This continuous treatment has allowed us to exploit short run effects as instrument.

6 Results

Table 5 shows the old-young elasticity parameter, $-\frac{1}{\lambda}$, by using different estimators. All regressions are weighted by the inverse of the wage gap variance to reduce the bias of the cells with small sample size[15]. In the first two columns we show the results by using ordinary least squares, OLS. As discussed in previous sections, the estimates are biased towards zero both without and with the time-occupation fixed effects. This finding is in line with the literature, that highlights the positive correlation between labor gap and *offsetting mechanism* in every region-occupation-year cell.
In the following six columns we use IV methodology by exploiting different estimators. From the third to the sixth column, we show the results by using two stage least squares, 2SLS, and the limited information maximum likelihood, LIML. As pointed out by Angrist and Krueger (1991), 2SLS and LIML estimates have to be very close in an overidentified framework because asymptotically they have the same distribution. The point estimates are -0.250 and -0.259 for 2SLS and LIML without occupation-time fixed effects and -0.168 and -0.171 for 2SLS and LIML including them. The quite similar results do not show any problem in the specification. Furthermore, the specifications pass the F-statistic and the overidentification tests. The former is 45.46 and 14.69, respectively, without and with occupation-time fixed effects and the other one cannot reject the null hypothesis of good specification at a significance level of 5%. The last two columns show the estimates with a continuously-updated GMM estimator, that allows for heteroskedastic and autocorre-

[14] P-values of Harris-Tzavalis unit-root test are: 1.00 for old foreign labor supply, .798 for young foreign labor supply, .1406 for old native labor supply and .9995 for young native labor supply

[15] There is a wide debate about the weight to be used. OP (2012) use the sample size in every specific cell as weight, while Borjas et.al (2012) in a comment to their paper say that is better to use the inverse of the wage variance.

lation disturbances, we add this estimation in order to take into account of a possible correlation among different shocks. Estimates confirm the previous results.
In Table 7 we show the estimates with employment cell weight to be sure that wrong weights drive our estimates. The estimates are quite similar, not showing any difference to use different weights. Results in Table 5 and 7 prove that there is not difference between the wage variance weight and employment-cell weight when the model is well specified.
In Table 8 we use a control function approach[16]. This approach used by Wooldridge (2015) is suitable for our aim, because residuals contain the endogeneity source that we cannot control in our model. In this way by adding this part in the main regression we control directly for the bias. Residuals' parameter captures the *offsetting mechanism* of shocks. Indeed, as showed in Table 8, the estimate has the opposite sign of our old-young elasticity parameter and more or less the same magnitude.
Hence, by using the control function approach we have not only the elasticity parameter but also the relative bias. In particular when we add the occupation-time fixed effects the parameter is even closer in absolute value to the elasticity estimate.
Old-young elastacity of substitution, λ, is between 4 and 6. The results are in line with Ottaviano and Peri (2012), Borjas (2003) and Card and Lemieux (2001) estimates when they use 8 level of experience and 4 level of education. Our result differ from Ottaviano and Peri (2012), when they use old and young as a proxy for experience. They find an elasticity of substitution around 3[17]. This result shed lights on time correlation bias.

7 Robustness checks

7.1 The comparison with migration instrument

Ottaviano and Peri (2012) and Borjas (2003) exploit foreign workers as instrument to identify age-group elasticity of substitution[18]. They assume that foreign labor supply shocks in the foreign labor force in each region-occupation-year cell, once added fixed effects, identify the elasticity of substitution between old and young workers.

In subsection 4.2 we find that the effect of a change in foreign labor supply on the

[16]In the control function approach you have to run an IV first stage, then get the residuals and put them in the main regression.
[17]Their estimate is equal to -0.31
[18]An other instrument widely used by researchers is the shift-share instrument (Altonij and Card, 1991) that exploits migrant enclaves in the previous decades to create an instrument exogenous with respect to current economic conditions. Unfortunately we cannot compare our methodology with that, because our dataset does not have information later than 1985 and because the level of inflows in decades before 1995 is almost null or is selected among high skilled workers (before 1990 in Italy there was not a labor migration policy, so for migrant was almost impossible to come in Italy with a labor VISA.)

labor gap is not constant since it depends on native labor supply changes. In order to take that into account, we consider both native and foreign labor supply shocks in each region-occupation-age-citizenship cell aiming at having an unbiased estimate.
In this subsection, we compare our specification with the one used by Borjas (2003) and Ottaviano and Peri (2012) by using their instrument to estimate the elasticity of substitution between old and young workers.

Table 9 shows estimates close to Ottaviano and Peri (2012)[19] when we do not control occupation-time fixed effects. The results change when we add them. The parameter is not more significant with standard error quite large. Our explanation is that large standard errors are due to the correlation between instrument and error term. The missing information on native labor supply changes in other skilled categories might create a correlation between the foreign instrument and the *offsetting mechanism* process.

In our setting we add both native and foreign labor supply changes, which help us to rule out any possible correlation with the long run adjustments.

7.2 Labor demand shocks

In our paper we exploit short run changes in every region-occupation-age-citizenship cell. We state that the short run changes are supply driven assuming that the possible labor demand shocks are captured by fixed effects. In this section we test this assumption by using an instrument that is based on labor demand shock.

We exploit the information on unemployment support mechanisms provided in our dataset. The Italian government helps firms to face crisis periods by paying part of employee salaries for a given period[20], when employers suspend temporary employment relationships. During the suspension firms cannot hire other workers in order to hold the benefit and, at the same time, employees cannot engage in another job. This tool has been created mainly to help the manufacturing sector, where most of the workers are engaged. We, thus, use the information whether a worker is in this program to capture a labor demand shock due to a firm's temporary crisis.

Our first stage is the labor gap on the log of the number of temporary suspended workers. The sample is reduced to 322 observations from 1996-2004 since we have some cells that do not experience any temporary suspension. Table 10 shows the results. The estimates are not significant and F-stat of the instrument is very low. The F stat is very low showing how a labor demand-driven instrument is not suitable to provide an exogenous variation to identify labor gap.

[19] Our estimates are little bit larger because we use a regional approach as opposed to a national one. For this reason the bias is smaller than Ottaviano and Peri (2012).

[20] The so called "Cassa Integrazione"

7.3 Relaxing time invariant assumption on the first-stage parameters

In subsection 4.2 we have assumed that the derivatives of each specific region-occupation-age labor input with respect to a change in both native and foreign employment is constant over time. In this subsection we show the results when derivatives vary over time.
In order to add a time dimension to parameters, we multiply the instruments by a linear time trend[21]. Rearranging Eq. (7) we get:

$$\Delta(ln(L_{rsOt}) - ln(L_{rsYt})) = \Sigma_c(\beta_{Oc}(1 + \mathit{trend})\frac{\Delta L_{rsOct}}{L_{rsOt}} - \beta_{Yc}(1 + \mathit{trend})\frac{\Delta L_{rsYct}}{L_{rsYt}}) \qquad c = N, F \quad (14)$$

Table 11 shows the results using this broader set of instruments that takes into account of time dimension. The results are quite similar to the time constant ones. Hence, without loss of generality, we can assume time invariant parameters.

8 Conclusions

How much old and young workers are substitutes in production is of great interest worldwide. Demographic changes have affected the composition of the labor force. This paper tries to shed light on the degree of substitutability between old and young workers in production.

In line with the tradition on this topic, we use a nested-CES framework at regional level to derive the elasticity of substitution between old and young workers. The choice of a local approach allows us to rule out any possible misleading effect due to huge differences across Italian regions. We can get a more precise estimate of the overall elasticity of substitution controlling for different local growth patterns.

In the literature, adding specific skill fixed effects is generally used to control for different demand shifts. Still, possible biases might remain due to log-run factors. In particular, skill-specific unobservable long-run adjustments offset any variation suited to study the effect of a skill-specific labor force change on the skill-specific wages biasing the elasticity estimates towards zero. Studying the dynamics of the *offsetting mechanism*, we identify the bias and create an instrument to overcome the omitted variable problem. We exploit the variation triggered by Italian reforms to create a set of instruments exogenous to skill-specific . The estimated elasticity of substitution is between 4 and 6, in line with previous findings.

[21] We try also to interact time dummies with every instrument. The results are not different from adding a linear trend.

In a global scenario, young workers concern about the longer working life of some old workers. Our results show old and young workers are imperfect substitution in production. Hence, policies aiming at reducing the youth unemployment should not consider the early retirement as a solution. Instead, they should look at specific characteristics of young workers to increase the match between firms needs and young worker skills.

References

[1] Altonij, Joseph G. and David Card, “The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives” in *Immigration, Trade and Labor Market*, John Abowd and Richard. Freeman,eds. (Chicago, University of Chicago Press:1991).

[2] Amiti, Mary, Oleg Itskhoki, Jozef Konings, “International Shocks, Variable Markups, and Domestic Prices”, *The Review of Economic Studies*, Feb 2019.

[3] Angrist, Joshua D. and Alan B. Krueger, “Does Compulsory School Attendance Affect Schooling and Earnings?”, *The Quarterly Journal of Economics*, Vol. 106, No. 4 (Nov., 1991), pp. 979-1014

[4] Arellano, Manuel and Bover, Olympia, (1995), “Another look at the instrumental variable estimation of error-components models”, *Journal of Econometrics*, 68, issue 1, p. 29-51.

[5] Barbagli, M., A. Colombo, G. Sciortino (a cura di), “I sommersi e i sanati. Le regolarizzazioni degli immigrati in Italia”, Il Mulino, Bologna 2004, pag. 51.

[6] Bertoni, Marco and Giorgio Brunello, “Does A Higher Retirement Age Reduce Youth Employment?,” Working paper 2017.

[7] Boeri, T., P. Garibaldi, and E. Moen (2017). “Closing the Retirement Door and the Lump of Labor” WorkINPS Paper.

[8] Borjas, G. J., (1999), “The economic analysis of immigration”, ch. 28, p. 1697-1760 in Ashenfelter, O. and Card, D. eds., *Handbook of Labor Economics*, vol. 3, Part A, Elsevier.

[9] Borjas,G. J., “The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market”, *Quarterly Journal of Economics*, 118 (2003), 1335-1374.

[10] Borjas, G. J., Grogger, J. and Hanson, G. H. (2012), “Comment: On Estimating Elasticities of Substition”, *Journal of the European Economic Association*, 10: 198-210.

[11] Bovini, G. and M. Paradisi, “Labor Substitutability and the Impact of Raising the Retirement Age,” Working paper 2018.

[12] Brugiavini, A., & Peracchi, F. (2008). “Youth Unemployment and Retirement of the Elderly: The Case of Italy”. Ssrn.

[13] David Card, “Chapter 30 - The Causal Effect of Education on Earnings”,Editor(s): Orley C. Ashenfelter, David Card, Handbook of Labor Economics, Elsevier, Volume 3, Part A, 1999, Pages 1801-1863, ISSN 1573-4463, ISBN 9780444501875.

[14] Card, David, and Thomas Lemieux “Can Falling Supply Explain the Rising Returns to College for Younger Men? A Cohort Based anaysis”, *Quarterly Journal of Economics*, 116 (2001), 705-746.

[15] D'Amuri, F., Gianmarco I.P. Ottaviano, Giovanni Peri, "The labor market impact of immigration in Western Germany in the 1990s", *European Economic Review*, Volume 54, Issue 4, 2010, Pages 550-570.

[16] Devillanova, C., Fasani, F., Frattini, T. (2018). "Employment of Undocumented Immigrants and the Prospect of Legal Status: Evidence from an Amnesty Program". ILR Review, 71(4), 853–881.

[17] Di Porto, Edoardo, Enrica Maria Martino and Paolo Naticchioni, 2018. "Back to Black? The Impact of Regularizing Migrant Workers," CSEF Working Papers 517, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.

[18] Freeman, Richard B., "The Effect of Demographic Factors on Age-Earnings Profiles", *Journal of Human Resources* XIV:3 (Summer 1979): pp 289-318.

[19] Gruber, Jonathan and David A. Wise, "Social Security Programs and Retirement around the World: The Relationship to Youth Employment", University of Chicago Press, 2010.

[20] Hamermesh, Daniel, "The Demand for Labor in the Long Run," in O. Ashenfelter and R. Layard, *Handbook of Labor Economics*, North-Holland, 1986.

[21] Jaeger, David A., Joakim Ruist, Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration" *NBER Working Paper* 24285 (2018)

[22] Katz, Lawrence F., and Kevin M. Murphy. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *The Quarterly Journal of Economics*, vol. 107, no. 1, 1992, pp. 35–78.

[23] Llull, J. (2018). "Immigration, wages, and education: A labour market equilibrium structural model". *Review of Economic Studies*, 85(3), 1852-1896.

[24] Manacorda, M. , Manning, A. and Wadsworth, J. (2012), "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain". *Journal of the European Economic Association*, 10: 120-151.

[25] Mohen, P. (2019). "The Impact of the Retirement Slowdown on the U.S. Youth Labor Market". Working Paper

[26] Munnell, A. H., & Wu, A. Y. (2013). "Will Delayed Retirement By the Baby Boomers Lead to higher unemployment among younger workers?".

[27] Nickell, Stephen, (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica*, 49, issue 6, p. 1417-26.

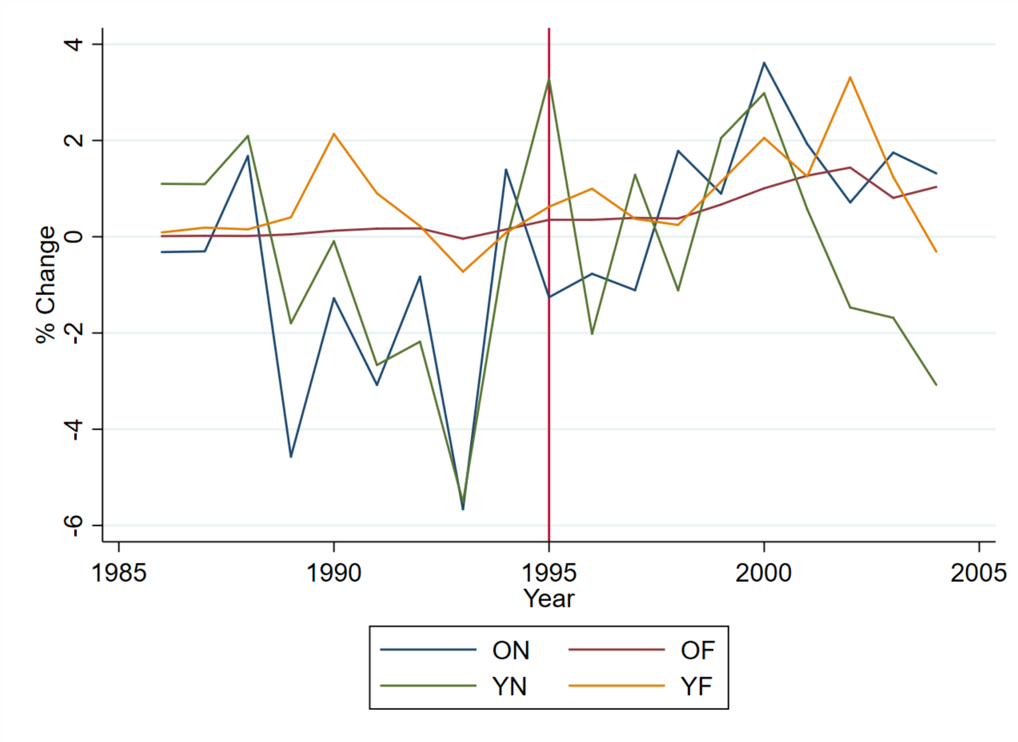
[28] Ottaviano, Gianmarco I. P., and Giovanni Peri, "Rethinking the Effect of Immigration on Wages" *Journal of the European Economic Association* 10.1 (2012): 152-197.

[29] Venturini, A., & Villosio, C. (2008). "Labour-market assimilation of foreign workers in Italy". Oxford Review of Economic Policy, 24(3), 518-542.

[30] Wooldridge, Jeffrey M., "Control Function Methods in Applied Econometrics," Journal of Human Resources 50, 420-445, March 2015.

Tables and Figures

Figure 1: Ratios of age-citizenship yearly employment changes to the relative yearly age employment between 1985 and 2004



ON: old natives. OF: old foreigners. YN:young native. YF: young foreigners.

Table 1: Blue collar subgroup shares within all sample, within the subgroup sample and within the sector sample

		Old			Young		
Macro Region	Place of birth	(1)	(2)	(3)	(1)	(2)	(3)
North	Foreign-born	1.695	5.122	19.616	6.947	10.384	80.384
	Native-born	15.880	47.982	32.917	32.362	48.371	67.083
Centre	Foreign-born	0.434	1.312	21.110	1.623	2.426	78.890
	Native-born	5.674	17.144	36.439	9.897	14.793	63.561
South and Islands	Foreign-born	0.177	0.534	22.971	0.593	0.886	77.029
	Native-born	9.235	27.905	37.365	15.481	23.140	62.635
Subgroup Obs			168,968			341,571	

Notes: (1) subgroup share in the row sector with respect to all sample;(2) subgroup share in the row sector with respect to subgroup sample; (3) subgroup share in the row sector with respect to row-sector sample. Sample from 1995 to 2004.

Table 2: White collar subgroup shares within all sample, within the subgroup sample and within the sector sample

		Old			Young		
Macro Region	Broad Experience	(1)	(2)	(3)	(1)	(2)	(3)
North	Foreign-born	0.289	0.709	35.850	0.517	0.873	64.150
	Native-born	23.499	57.635	38.952	36.830	62.185	61.048
Centre	Foreign-born	0.172	0.422	46.964	0.195	0.328	53.036
	Native-born	8.667	21.256	42.406	11.771	19.874	57.594
South and Islands	Foreign-born	0.088	0.215	39.955	0.132	0.222	60.045
	Native-born	8.058	19.762	45.164	9.783	16.518	54.836
Subgroup Obs			82,375			119,660	

Notes: (1) subgroup share in the row sector with respect to all sample;(2) subgroup share in the row sector with respect to subgroup sample; (3) subgroup share in the row sector with respect to row-sector sample. Sample from 1995 to 2004.

Table 3: Average log daily wage for male blue collar workers, 1995-2004

Macro Region	Broad Experience	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
North	Old	4.4465	4.4394	4.4551	4.4504	4.4455	4.4258	4.4259	4.4154	4.4040	4.4255
	Young	4.2900	4.2879	4.3038	4.3111	4.3045	4.2988	4.2974	4.2924	4.2915	4.3116
Centre	Old	4.4341	4.4239	4.4318	4.4214	4.4174	4.4127	4.3780	4.3712	4.3611	4.3762
	Young	4.2591	4.2570	4.2639	4.2703	4.2615	4.2619	4.2587	4.2425	4.2473	4.2639
South and Islands	Old	4.4278	4.4087	4.4175	4.3898	4.3941	4.3781	4.3700	4.3526	4.3514	4.3596
	Young	4.2612	4.2430	4.2471	4.2269	4.2184	4.2410	4.2381	4.2430	4.2397	4.2609

Notes: The table reports the mean of the log daily wage of workers in each region-age group. All wages are deflated to 2001 euro using the Italian CPI index from OECD database.

Table 4: Average log daily wage for male white collar workers, 1995-2004

Macro Region	Broad Experience	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
North	Old	4.9591	4.9633	4.9865	4.9917	5.0035	5.0045	4.9935	4.9927	4.9949	5.0029
	Young	4.6578	4.6478	4.6752	4.6835	4.6943	4.6969	4.6991	4.7036	4.6916	4.7041
Centre	Old	4.9576	4.9545	4.9712	4.9756	4.9598	4.9522	4.9507	4.9392	4.9479	4.9503
	Young	4.6336	4.6288	4.6339	4.6441	4.6497	4.6499	4.6426	4.6455	4.6350	4.6320
South and Islands	Old	4.8804	4.8644	4.8806	4.8623	4.8648	4.8614	4.8413	4.8333	4.8241	4.8364
	Young	4.5260	4.5136	4.5329	4.5257	4.5026	4.4920	4.4916	4.4737	4.4747	4.4880

Notes: The table reports the mean of the log daily wage of workers in each region-age group. All wages are deflated to 2001 euro using the Italian CPI index from OECD database.

Table 5: Estimated old-young elasticity of substitution, $-\frac{1}{\lambda}$, weighted by the inverse of dependent variable variance

$ln(\frac{w_{rsOt}}{w_{rsYt}})$	OLS (1)	OLS (2)	LIML (3)	LIML (4)	2SLS (5)	2SLS (6)	GMM-CUE (7)	GMM-CUE (8)
$ln(\frac{L_{rsOt}}{L_{rsYt}})$	-0.0893***	-0.0157	-0.259***	-0.171***	-0.250***	-0.168***	-0.215***	-0.164***
	(0.0322)	(0.0366)	(0.0518)	(0.0510)	(0.0494)	(0.0499)	(0.0461)	(0.0484)
Skill x Time FE	NO	YES	NO	YES	NO	YES	NO	YES
P-value Lm-stat			0.000298	0.000745	0.000298	0.000745	0.000298	0.000745
F-stat			45.46	14.69	45.46	14.69	45.46	14.69
P-value J-stat			0.0773	0.389	0.0742	0.386	0.0742	0.386
Observations	342	342	342	342	342	342	342	342

Notes: All regressions include region-year fixed effects and region-occupation fixed effects. The reported standard errors are clustered by region-occupation. The regressions are weighted by the inverse of the wage ratio variance. The dependent variable is the log ratio of average old and young worker wages in a region-occupation-age-time-specific cell and the explanatory variable is the log of the ratio between old and young employment inside the same cell. LM test is for underidentification test, F statistic is for weak identification test (Cragg-Donald or Kleibergen-Paap) and J test is for overidentification test. Time span covers from 1996-2004. * p<0.10, ** p<0.05, *** p<0.01

Table 6: First stage and Reduced Form weighted by the inverse of dependent variable variance

	First stage		Reduced Form	
$\frac{\Delta L_{rsOFt}}{L_{rsOt-1}}$	1.399	-0.0254	-0.659***	-0.138
	(0.907)	(0.716)	(0.213)	(0.253)
$\frac{\Delta L_{rsYFt}}{L_{rsYt-1}}$	-0.448*	-0.160	0.102	0.0300
	(0.246)	(0.521)	(0.105)	(0.183)
$\frac{\Delta L_{rsONt}}{L_{rsOt-1}}$	0.433***	0.508***	-0.0850**	-0.0658*
	(0.119)	(0.0918)	(0.0391)	(0.0393)
$\frac{\Delta L_{rsYNt}}{L_{rsYt-1}}$	-0.743***	-0.569***	0.166***	0.115**
	(0.107)	(0.105)	(0.0520)	(0.0538)
Skill x Time FE	NO	YES	NO	YES
Observations	342	342	342	342

Notes: All regressions include region-year fixed effects and region-occupation fixed effects. The reported standard errors are clustered by region-occupation. The regressions are weighted by the inverse of wage ratio variance. The explanatory variables are measured as the yearly change in the old (young) native (foreign) group over the total old (young) employment. Time span covers from 1996-2004. * p<0.10, ** p<0.05, *** p<0.01

Table 7: Estimated old-young elasticity of substitution, $-\frac{1}{\lambda}$, weighted by the size of independent variable

$ln(\frac{w_{rsOt}}{w_{rsYt}})$	OLS (1)	OLS (2)	LIML (3)	LIML (4)	2SLS (5)	2SLS (6)	GMM-CUE (7)	GMM-CUE (8)
$ln(\frac{L_{rsOt}}{L_{rsYt}})$	-0.0897***	-0.0168	-0.258***	-0.167***	-0.249***	-0.164***	-0.210***	-0.158***
	(0.0306)	(0.0343)	(0.0499)	(0.0492)	(0.0472)	(0.0480)	(0.0436)	(0.0470)
Skill x Time FE	NO	YES	NO	YES	NO	YES	NO	YES
P-value Lm-stat			0.000108	0.000190	0.000108	0.000190	0.000108	0.000190
F-stat			51.22	18.00	51.22	18.00	51.22	18.00
P-value J-stat			0.0525	0.402	0.0490	0.400	0.0490	0.400
Observations	342	342	342	342	342	342	342	342

Notes: All regressions include region-year fixed effects and region-occupation fixed effects. The reported standard errors are clustered by region-occupation. The regressions are weighted by the size of every region-occupation-time cell. The dependent variable is the log ratio of average old and young worker wages in a region-occupation-age-time-specific cell and the explanatory variable is the log of the ratio between old and young employment inside the same cell. LM test is for underidentification test, F statistic is for weak identification test (Cragg-Donald or Kleibergen-Paap) and J test is for overidentification test. Time span covers from 1996-2004. * p<0.10, ** p<0.05, *** p<0.01

Table 8: Control function estimates of elasticity of substitution $\frac{1}{\lambda}$

$ln(\frac{w_{rsOt}}{w_{rsYt}})$	Control Function (1)	(2)
$ln(\frac{L_{rsOt}}{L_{rsYt}})$	-0.250***	-0.168***
	(0.0435)	(0.0509)
Residuals	0.209***	0.180***
	(0.0559)	(0.0568)
Skill x Time FE	NO	YES
Observations	342	342

Notes: All regressions include region-year fixed effects and region-skill fixed effects. The reported standard errors are clustered by region-skill. The regressions are weighted by the inverse of the wage ratio variance. Time span covers from 1996-2004. * p<0.10, ** p<0.05, *** p<0.01

Table 9: Estimated old-young elasticity of substitution, $-\frac{1}{\lambda}$ by exploiting foreign employment as instrument

$ln(\frac{w_{rsOt}}{w_{rsYt}})$	OLS (1)	OLS (2)	2SLS (3)	2SLS (4)
$ln(\frac{L_{rsOt}}{L_{rsYt}})$	-0.0893*** (0.0322)	-0.0157 (0.0366)	-0.298*** (0.104)	-0.0167 (0.338)
Skill x Time FE	NO	YES	NO	YES
P-value Lm-stat			0.00981	0.415
F-stat			5.285	0.336
Observations	342	342	334	334

Notes: All regressions include region-year fixed effects and region-occupation fixed effects. The reported standard errors are clustered by region-occupation. The dependent variable is the log ratio of average old and young worker wages in a region-occupation-age-time-specific cell and the explanatory variable is the log of the ratio between old and young employment inside the same cell. The instrument is the log oh the total region-occupation-time specific foreign employment. LM test is for underidentification test and F statistic is for weak identification test (Cragg-Donald or Kleibergen-Paap). Time span covers from 1996-2004. * p<0.10, ** p<0.05, *** p<0.01

Table 10: Estimated old-young elasticity of substitution, $-\frac{1}{\lambda}$, by exploiting temporary laid-off workers as instrument

$ln(\frac{w_{rsOt}}{w_{rsYt}})$	OLS (1)	OLS (2)	2SLS (3)	2SLS (4)
$ln(\frac{L_{rsOt}}{L_{rsYt}})$	-0.0915***	-0.0264	-0.268	0.0263
	(0.0310)	(0.0535)	(0.220)	(0.796)
Skill x Time FE	NO	YES	NO	YES
P-value Lm-stat			0.100	0.619
F-stat			1.055	0.106
Observations	342	342	322	322

Notes: All regressions include region-year fixed effects and region-occupation fixed effects. The reported standard errors are clustered by region-occupation. The regressions are weighted by the inverse of the wage ratio variance. The regression are run only on the manufacturing sector. The dependent variable is the log ratio of average old and young worker wages in a region-occupation-age-time-specific cell and the explanatory variable is the log of the ratio between old and young employment inside the same cell. The instrument is the log of the total number of temporary laid-off workers. LM test is for underidentification test and F statistic is for weak identification test (Cragg-Donald or Kleibergen-Paap). Time span covers from 1995-2004. * p<0.10, ** p<0.05, *** p<0.01

Table 11: Estimated old-young elasticity of substitution, $-\frac{1}{\lambda}$, by adding a linear trend to instruments

$ln(\frac{w_{rsOt}}{w_{rsYt}})$	OLS (1)	OLS (2)	LIML (3)	LIML (4)	2SLS (5)	2SLS (6)	GMM-CUE (7)	GMM-CUE (8)
$ln(\frac{L_{rsOt}}{L_{rsYt}})$	-0.0893***	-0.0157	-0.268***	-0.168***	-0.249***	-0.164***	-0.189***	-0.144***
	(0.0322)	(0.0366)	(0.0532)	(0.0556)	(0.0476)	(0.0541)	(0.0353)	(0.0406)
Skill x Time FE	NO	YES	NO	YES	NO	YES	NO	YES
P-value Lm-stat			0.00321	0.00808	0.00321	0.00808	0.00321	0.00808
F-stat			46.73	12.09	46.73	12.09	46.73	12.09
P-value J-stat			0.208	0.802	0.192	0.798	0.192	0.798
Observations	342	342	342	342	342	342	342	342

Notes: All regressions include region-year fixed effects and region-occupation fixed effects. The reported standard errors are clustered by region-occupation. The dependent variable is the log ratio of average old and young worker wages in a region-occupation-age-time-specific cell and the explanatory variable is the log of the ratio between old and young employment inside the same cell. LM test is for underidentification test, F statistic is for weak identification test (Cragg-Donald or Kleibergen-Paap) and J test is for overidentification test. Time span covers from 1995-2004. * p<0.10, ** p<0.05, *** p<0.01

Appendix A

By defining labor gap for old and young workers as function of native and foreigners, we get

$$ln(L_{rsOt}) - ln(L_{rsYt}) = ln(f(L_{rsONt}, L_{rsOFt})) - ln(f(L_{rsYNt}, L_{rsYFt})) \tag{15}$$

Computing the differential, we get:

$$d(ln(L_{rsOt}) - ln(L_{rsYt})) = d(ln(f(L_{rsONt}, L_{rsOFt})) - ln(f(L_{rsYNt}, L_{rsYFt}))) =$$

$$= \frac{1}{L_{rsOt}} \frac{\partial L_{rsOt}}{\partial L_{rsONt}} dL_{rsONt} + \frac{1}{L_{rsOt}} \frac{\partial L_{rsOt}}{\partial L_{rsOFt}} dL_{rsOFt} - (\frac{1}{L_{rsYt}} \frac{\partial L_{rsYt}}{\partial L_{rsYNt}} dL_{rsYNt} +$$

$$+ \frac{1}{L_{rsYt}} \frac{\partial L_{rsYt}}{\partial L_{rsYFt}} dL_{rsYFt}) \tag{16}$$

Labor ratio differential in discrete is:

$$\Delta(ln(L_{rsOt}) - ln(L_{rsYt})) = \Sigma_c(\beta_{Oc} \frac{\Delta L_{rsOct}}{L_{rsOt}} - \beta_{Yc} \frac{\Delta L_{rsYct}}{L_{rsYt}}) \qquad c = N, F \tag{17}$$

where

$$\beta_{Oc} = \frac{\partial L_{rsOt}}{\partial L_{rsOct}} \qquad \beta_{Yc} = \frac{\partial L_{rsYt}}{\partial L_{rsYct}} \tag{18}$$

with c indicating if they are natives or foreigners.

Appendix B

By assuming that the residuals, x, obtained by applying Frisch–Waugh–Lovell theorem to labor gap first difference, follow an AR(1) process:

$$x_{rst} = \rho x_{rst-1} + \varepsilon_{rst} \tag{19}$$

By subtracting x_{rst-1} on both sides we obtain:

$$\Delta x_{rst} = (\rho - 1)x_{rst-1} + \varepsilon_{rst} \tag{20}$$

By substituting the first difference in the labor demand we have:

$$ln(\frac{w_{rsOt}}{w_{rsYt}}) = \beta \Delta x_{rst} + u_{rst} \tag{21}$$

Where y_{rst} is the residuals from wages by using the Frisch–Waugh–Lovell theorem and t goes from 2 to T. We omit to multiply the parameter from the first stage regression.

$$\operatorname*{plim}_{N\to\infty} \widehat{\beta} = \beta + \frac{\text{plim}_{N\to\infty} \frac{1}{N}\Sigma_{i=1}^{N}(x_{rst} - x_{rs.})(u_{rst} - u_{rs.})}{\text{plim}_{N\to\infty} \frac{1}{N}\Sigma_{i=1}^{N}(x_{rst} - x_{rs.})^2} \tag{22}$$

$$\operatorname*{plim}_{N\to\infty} \widehat{\beta} - \beta = \frac{\text{plim}_{N\to\infty} \frac{1}{N}\Sigma_{i=1}^{N}(x_{rst} - x_{rs.})(u_{rst} - u_{rs.})}{\text{plim}_{N\to\infty} \frac{1}{N}\Sigma_{i=1}^{N}(x_{rst} - x_{rs.})^2} = \frac{A}{B} \tag{23}$$

$$A = \underbrace{E[x_{rst}u_{rst}]}_{\text{endogeneity bias}} \underbrace{-E[x_{rst}u_{rs.}] - E[x_{rs.}u_{rst}] + E[x_{rs.}u_{rs.}]}_{\text{Nickell bias}} \tag{24}$$

$$B = E[x_{rst}^2] - 2E[x_{rst}x_{rs.}] + E[x_{rs.}^2] \tag{25}$$

Let u_{rst} equals to the sum of a random effect and a *offsetting mechanism*:

$$u_{rst} = \xi_{rst} + f(\varepsilon_{rst-1}) \tag{26}$$

Let $f(\epsilon_{rst-1})$ linear and function of labor supply shock at t-1:

$$f(\epsilon_{rst-1}) = \varepsilon_{rst-1} \tag{27}$$

Let ρ is equal to one:

$$\Delta x_{rst} = \varepsilon_{rst} \tag{28}$$

Then A turns into:

$$A = E[\varepsilon_{rst}\varepsilon_{rst-1}] - E[\varepsilon_{rst}\varepsilon_{rs.-1}] - E[\varepsilon_{rs.}\varepsilon_{rst-1}] + E[\varepsilon_{rs.}\varepsilon_{rs.-1}] \tag{29}$$

Showing the time means of A:

$$A = E[\varepsilon_{rst}\varepsilon_{rst-1}] - \tfrac{1}{T-1}E[\varepsilon_{rst}\textstyle\sum_{j=2}^{T}\varepsilon_{rsj-1}] - \tfrac{1}{T-1}E[\textstyle\sum_{j=2}^{T}\varepsilon_{rsj}\varepsilon_{rst-1}] +$$

$$\frac{1}{(T-1)^2}E[\sum_{j=2}^{T}\varepsilon_{rsj}\sum_{j=2}^{T}\varepsilon_{rsj-1}] \tag{30}$$

Solving covariates:

$$A = 0 - \frac{1}{T-1}\sigma_\varepsilon^2 - \frac{1}{T-1}\sigma_\varepsilon^2 + \frac{1}{(T-1)^2}(T-2)\sigma_\varepsilon^2 \tag{31}$$

$$A = -\frac{T}{(T-1)^2}\sigma_\varepsilon^2 \tag{32}$$

Doing the same for B:

$$B = E[\varepsilon_{rst}^2] - 2E[\varepsilon_{rst}\varepsilon_{rs.}] + E[\varepsilon_{rs.}^2] \tag{33}$$

$$B = E[\varepsilon_{rst}^2] - \frac{2}{(T-1)}E[\varepsilon_{rst}\sum_{j=2}^{T}\varepsilon_{rsj}] + \frac{1}{(T-1)^2}E[\sum_{j=2}^{T}\varepsilon_{rsj}^2] \tag{34}$$

$$B = \sigma_\varepsilon^2 - \frac{2}{T-1}\sigma_\varepsilon^2 + \frac{1}{(T-1)^2}(T-1)\sigma_\varepsilon^2 \tag{35}$$

$$B = \sigma_\varepsilon^2\frac{T-2}{T-1} \tag{36}$$

Computing the ratio between A and B:

$$\frac{A}{B} = -\frac{T}{(T-2)(T-1)} \tag{37}$$

Household gender disparity and portfolio choices

Giulia Bettin[1] Emma Sarno[2] Md Taslim Uddin[3]
Alberto Zazzaro[4]

May 22, 2024

[1]Università Politecnica delle Marche and MoFiR
[2]Università di Napoli L'Orientale
[3]Asian University for Women
[4]Università di Napoli Federico II, CSEF and MoFiR

Abstract

This paper examines how women's economic empowerment within households affects financial portfolio choices. Using data from the Bank of Italy's Survey on Household Income and Wealth for 2004–2020, we measure empowerment via the female income share and examine its association with participation in risky assets and the share of financial wealth invested in equities and private bonds. The results show that a higher female income share is, on average, associated with lower exposure to risky assets. However, this aggregate effect conceals a key asymmetry: in male-headed households, a greater female income share reduces risky investment, whereas in female-headed households it increases it. Tests based on imputed individual risk preferences provide little support for a bargaining model in which income shares reweight partners' risk attitudes. Instead, the evidence supports a unitary decision-making framework à la Becker, in which the identity and confidence of the household financial decision-maker are central. Financial-sector employment of the household head substantially attenuates these effects, whereas similarity in partners' risk preferences does not, pointing to self-confidence rather than altruistic preference aggregation as the primary mechanism underlying household portfolio choices.

Keywords: Household finance, Gender disparity, Asset allocation, Household unitary decision-making, Household bargaining, Risk aversion.

JEL Codes: D13; D14; G11; J16.

1 Introduction

A large body of research in household finance and behavioral economics documents systematic gender differences in financial behavior, linking these differences to variations in risk aversion, self-confidence, financial literacy, expectations, and gender norms (Barber and Odean, 2001; Lusardi and Mitchell, 2008; Croson and Gneezy, 2009; Ke, 2021; Guiso and Zaccaria, 2023). These findings are central to our understanding of individual portfolio choice. Yet most financial decisions are not made by isolated individuals, but within households, where preferences, resources, information, and decision rights interact.

This paper asks whether women's economic empowerment within the household affects household financial investment decisions. More specifically, we study whether the share of household income earned by women and men affects both the probability that the household holds risky assets and the fraction of financial wealth invested in such assets. We also investigate the mechanism by which this effect operates. Does the relative contribution to household income affect portfolio choices by shifting bargaining power between spouses or partners? Or, are financial decisions delegated to a single household member (male or female)? In this case, does the relative contribution of women and men to household income affect investment choices through its impact on the decision-maker's self-confidence, or through its influence on the degree of altruism with which the preferences of the partner are taken into account?

Addressing these questions is important because household financial decisions differ fundamentally from individual ones. When individuals form a household, investment choices are no longer determined by a single set of preferences, but instead reflect intra-household interactions shaped by marital status, differences in economic autonomy, and the allocation of decision-making power between partners (Love, 2010; ?; Bertocchi et al., 2014; Gomes et al., 2021). In this context, a partner's contribution to household income may play a dual role. First, it can shift the balance of decision-making power within the household, thereby giving greater weight to his/her preferences and greater responsibility in household financial decisions. Second, it may affect the behavior of those who effectively manage the household portfolio, for instance, by increasing or reducing their confidence in their own financial skills and decisions, or by altering the extent to which they accommodate the partner's preferences.

The distinction is theoretically relevant and empirically consequential. A bargaining view of the household implies that financial choices result from a weighted aggregation of individual preferences, with weights depending on each partner's economic resources or outside options (McElroy and Horney, 1981). Under this approach, the effect of women's income share on risky investment should depend on the relative risk aversion of women and men within the household. If women are more risk-averse than men, a higher female income share should reduce risky investment; if they are less risk-averse, it should increase it; and if risk preferences are similar, it should have little or no effect. Importantly, these effects should hold regardless of whether the financial decision-maker is male or female. By contrast, a unitary or dictatorship view of the household emphasizes who controls financial decisions (Becker, 1981b). In this case, the same increase in women's income share may have opposite effects depending on whether the household portfolio is managed by a woman or by a man.

We study these issues using the Bank of Italy's Survey on Household Income and Wealth (SHIW) over the period 2004–2020. The SHIW is particularly well-suited for our purposes because it provides detailed information on household income, wealth, demographic characteristics, financial assets, individual income and education, and the identity of the household member responsible for financial decisions, as well as her/his financial literacy and risk propensity. We measure women's economic empowerment using the female income share, defined as the ratio of total female income to total household income. Our main outcomes are two standard measures of risky financial investment: an indicator for holding risky assets and the share of financial wealth invested in risky assets. In the empirical analysis, risky assets are defined as the sum of equities and private bonds.

Before turning to mixed-gender households, the Italian data provide a useful benchmark on the individual gender gap in portfolio choice. Among single-person households, women invest substantially less in risky assets than men (see Table 1). Female singles hold equities in only 1.91 percent of cases, compared with 6.32 percent for male singles. The difference is even more striking in portfolio shares: equities account for 0.09% of financial wealth among single women and 2.7% among single men. A similar pattern emerges for the broader risky-asset measure, with holding rates of 7.84 percent for single women and 13.92 percent for single men, and portfolio shares of 5.99 and 9.65 percent, respectively. These descriptive figures confirm that the standard gender gap documented in the international literature is clearly present in the Italian SHIW data.

The relevance of the Italian setting for studying the role of women's economic empowerment in household financial decisions is supported by the regression results reported in Table 2,

Table 1: Portfolio composition by gender

	Equity		Risky assets		
	Holding	*Share*	*Holding*	*Share*	Obs.
Female	1.91%	0.09%	7.84%	5.99%	10,426
Male	6.32%	2.7%	13.92%	9.65%	6,011

Source. Bank of Italy, SHIW (2004-2020). *Female* and *Male* denote single-person female and male households, respectively. Risky assets are defined as the sum of equities and private bonds. *Holding* is a dummy variable equal to 1 if the household holds equities or, alternatively, risky assets. *Share* denotes the proportion of the household's financial wealth invested in equities or, alternatively, in risky assets.

which are based on the full SHIW sample and control for household income and time effects. First, both pooled OLS estimates of asset holdings and pooled Tobit estimates of portfolio shares indicate that single-person households exhibit portfolio choices that differ significantly from those of multi-person households and are, on average, more risk-taking. Second, this effect is far from homogeneous across genders. The coefficient associated with being a single male is systematically larger than that associated with being a single female. The probability that a single male holds equities is 4.3 percentage points higher than for multi-person households, whereas for a single female it is only 1.3 percentage points higher. For the broader category of risky assets, the corresponding figures are 6.8 and 3.6 percentage points for single men and single women, respectively. The same gender gap is evident in portfolio shares, which are significantly higher for men than for women, both for equities and for risky assets more broadly.

These preliminary, descriptive results serve two purposes. First, they validate in our empirical setting the basic premise that women and men differ in their financial investment behavior. Second, they show that household composition matters for portfolio choice. The fact that single men and single women behave differently suggests that individual gender traits, preferences, confidence, or social norms are relevant. At the same time, because our main analysis focuses on mixed-gender households, it is necessary to examine how these individual characteristics are combined, weighted, or overridden when financial decisions are made jointly or delegated to a single household member.

Our empirical analysis proceeds in three steps. First, we estimate the average relationship between the female income share and household risky investment in the full sample of mixed-

Table 2: Portfolio choice by household composition

	Equity			
	Holding		*Share*	
Single	0.024***		0.170***	
	(0.003)		(0.020)	
Single female		0.013***		0.037*
		(0.002)		(0.022)
Single male		0.043***		0.287***
		(0.004)		(0.020)
Observations	61,508	61,508	49,830	49,830
	Risky assets			
	Holding		*Share*	
Single	0.048***		0.235***	
	(0.004)		(0.019)	
Single female		0.036***		0.175***
		(0.004)		(0.024)
Single male		0.068***		0.306***
		(0.006)		(0.025)
Observations	61,508	61,508	49,830	49,830

Source: Bank of Italy, SHIW (2004–2020). The table reports pooled OLS estimates for *Holding*, a dummy variable equal to 1 if the household holds equities or risky assets, and pooled Tobit estimates for *Share*, defined as the proportion of the household's financial wealth invested in equities or in risky assets. The baseline specification is $y_{it} = \alpha + \beta\, Single_{it} + \gamma\, \log(income_{it}) + \lambda_t + \epsilon_{it}$, where *Single* is a dummy equal to 1 if the household is single-person and *income* denotes household income. The extended specification distinguishes single-person households by the head of household's gender.

gender households. We use pooled linear probability models for risky asset holding and Tobit models for risky asset shares. The specifications control for household income, wealth, demographic characteristics, employment status, risk attitudes, marital status, homeownership, and region-by-year fixed effects. The baseline results show that a higher female income share is associated with lower investment in risky assets. This finding indicates that women's contribution to household income is not neutral for household portfolio allocation.

Second, we test whether this average effect is consistent with a bargaining model of household financial decision-making. To do so, we examine whether the effect of female income share varies with the relative risk aversion of women and men within the household. Since the SHIW directly observes risk attitudes only for the household head, we impute risk propensity for all adult household members using a random forest procedure. This allows us to classify households according to whether adult women are, on average, more risk-averse than men, less risk-averse than men, or equally risk-averse. The results do not support the bargaining interpretation. Female income share remains negatively related to risky investment even when women and men in the household display similar risk aversion. This is difficult to reconcile with a framework in which the only channel is the reweighting of male and female risk preferences.

Third, we examine whether the evidence is instead consistent with a unitary model in which the financial decision-maker's identity is central. We distinguish households by whether the person in charge of financial decisions is male or female and estimate selection-corrected models of household headship. The results reveal a striking asymmetry. In male-headed households, a higher female income share reduces both the probability of holding risky assets and the share of risky assets. In female-headed households, the effect is positive. This pattern is consistent with the idea that women's income share changes the self-confidence, authority, or perceived competence of the household financial decision-maker, rather than simply shifting the average weight of male and female preferences.

We then investigate the mechanism behind this unitary pattern. We consider two channels. The first is altruism: as the household head's contribution to household income increases, they may assign greater or lesser weight to the partner's preferences, depending on whether altruistic or self-interested motives prevail. If this were the main channel, the effect of female income share should disappear when women and men have similar risk preferences. The data do not support this prediction. The second mechanism is self-confidence in asset evaluation. In this interpretation, income contributions affect the decision-maker's confidence in assessing financial opportunities and risks. A higher female income share may strengthen confidence in female-headed households and weaken confidence or authority among men in male-headed households. Consistent with this view, the effect of female income share is attenuated when the household head is financially sophisticated, proxied by employment in the financial sector.

The paper contributes to the household finance literature in three ways. First, it connects the literature on gender differences in financial behavior to that on intra-household decision-making. We show that women's economic contribution within the household is a key determinant of risky investment, but that its effect cannot be understood without considering who manages household finances.

Second, the paper provides a direct empirical test of competing models of household portfolio choice. While bargaining models predict that the effect of women's income share should depend on relative preferences, our evidence points to the importance of decision rights and headship. This finding is closely related to evidence that intra-household financial decision-making is shaped by the identity of the person who controls resources and makes financial choices (Bertocchi et al., 2014; Ke, 2021). It also suggests that income shares influence household finance not only through formal bargaining power, but also through psychological and organizational channels.

Third, the paper identifies self-confidence as a relevant mechanism through which women's economic empowerment affects financial investment. This channel is consistent with the broader behavioral finance literature emphasizing overconfidence, underconfidence, and gender differences in financial decision-making (Barber and Odean, 2001; Lusardi and Mitchell, 2008). Our results suggest that confidence is not only an individual trait, but may also be shaped within the household by relative income contributions and by the gender of the household financial decision-maker.

Overall, our results show that women's income share matters for household portfolio allocation, but not in the simple way predicted by a standard bargaining model. Women's economic empowerment reduces risky investment on average, yet this average masks opposite effects across male- and female-headed households. The findings, therefore, highlight the importance of opening the black box of household decision-making. In household finance, who earns income matters; but who controls financial decisions matters just as much.

The remainder of the paper proceeds as follows. Section 2 places the paper within the related literature on gender differences in financial behavior, household finance, intra-household allocation, and social norms. Section 3 presents the SHIW data, the construction of the key variables, and the baseline empirical framework. Section 4 develops the theoretical framework and derives the empirical predictions of bargaining and unitary models of household portfolio choice, and presents the corresponding empirical evidence. Section 5 examines the mechanisms underlying the main results, distinguishing between altruism and self-confidence, and provides empirical evidence. Section 6 concludes.

2 Literature review

Four strands of literature are most closely related to this paper: gender differences in financial behavior; the effects of family structure on portfolio choice; models of intra-household portfolio allocation and decision-making; and the role of social norms, confidence, and psychological factors in household financial choices.

2.1 Gender differences in financial behavior

A large body of literature documents systematic gender differences in financial behavior and portfolio allocation. Compared with men, women are generally less likely to participate in financial markets, invest a smaller share of their wealth in risky assets, and adopt more conservative investment strategies (Jianakoplos and Bernasek, 1998; Sunden and Surette, 1998). Subsequent research links these differences to overconfidence, with Barber and Odean (2001) showing that men trade more aggressively and ultimately obtain lower net returns, as well as to lower financial literacy, more cautious return expectations, and reduced willingness to engage in risky financial environments (Dominitz and Manski, 2007; Lusardi and Mitchell, 2008; Croson and Gneezy, 2009; Charness and Gneezy, 2012). At the household level, however, these individual-level differences interact with bargaining dynamics, resource allocation, and decision authority in ways that individual-level analyses cannot fully capture.

2.2 Household finance: Family structure and portfolio choice

This insight motivates the growing literature on household finance, which emphasizes that portfolio allocation depends not only on individual preferences and characteristics, but also on marital status, family structure, and intra-household interactions (Gomes et al., 2021). Love (2010) shows that marital status transitions generate large and asymmetric portfolio adjustments by gender: divorce pushes men toward riskier allocations while moving women toward safer ones, with children further moderating these effects. Bertocchi et al. (2011) argue that marriage acts as an insurance mechanism for women, reducing background income risk and increasing participation in risky financial markets, though this effect weakens as female labor market participation rises.

2.3 Intra-household decision-making: Bargaining vs unitary models

A central issue concerns how household preferences are aggregated into financial decisions. The standard unitary model (Becker, 1981b) treats the household as a single decision-making unit maximizing a common utility function subject to pooled resources, while bargaining models (McElroy and Horney, 1981) predict that portfolio allocation shifts with changes in spouses' relative resources and outside options. Jianakoplos and Bernasek (2008) test these alternative frameworks directly and find that neither relative earnings nor other proxies for bargaining power significantly affect the risky asset share, a result more consistent with unitary decision-making, though the authors note that this is difficult to interpret without individual-level measures of risk aversion. Arano et al. (2010) qualify this finding: gender differences in retirement asset allocation disappear at the individual level after standard controls, but prove significant in the context of joint household decision-making. More structural approaches confirm that both partners' preferences matter, but in asymmetric ways. Addoum et al. (2016) and Addoum (2017) show that shocks to spouses' relative income generate changes in household risk aversion and risky asset holdings, with effects particularly pronounced in households where wives are more risk averse than husbands. Yilmazer and Lich (2015) document a starker asymmetry: only the risk tolerance of the spouse with greater bargaining power influences portfolio outcomes, while that of the less powerful partner is largely irrelevant. Gu et al. (2024) structurally estimate the distribution of bargaining power from household portfolio data across three countries, finding a significant gender gap that is only partially explained by observable characteristics such as income and employment, with the residual linked to perceived gender norms. Maura (2022) similarly finds that the collective model fits portfolio data substantially better than the unitary benchmark, and that household risk preferences shape the optimal risky share conditional on stock market participation.

2.4 Decision authority, financial headship, and social norms

An alternative perspective emphasizes who effectively controls financial decisions rather than how preferences are aggregated. Bertocchi et al. (2014), using the same SHIW data employed in this paper, show that the wife's probability of assuming financial responsibility rises with her relative age, education, and income, consistent with bargaining, but falls when she is employed, suggesting a specialization pattern that assigns management to the spouse with more available time. These patterns, however, are not fully explained by economic characteristics alone. A growing body of evidence points to gender norms as an independent determinant of who controls financial decisions and with what consequences. Guiso and Zaccaria (2023), also using SHIW data, exploit variation across Italian cohorts and regions in the gender of the household head to show that more egalitarian norms raise market participation, equity holdings, diversification, and returns, with patriarchal norms that began receding in the early 1990s, when a pension reform raised the cost of adhering to traditional gender roles, a shift directly relevant to our sample period. Ke (2021), using data from U.S. household surveys, shows that the financial knowledge of the husband translates into stock market participation more readily than equivalent knowledge held by the wife, a pattern traced to gender identity norms that suppress women's effective voice in financial decisions, as confirmed by a randomized experiment.

3 Data and Baseline Model

3.1 Data Sources

We utilize data from the Bank of Italy's Survey on Household Income and Wealth (SHIW), which is conducted every other year on a representative sample of the Italian resident population. The SHIW provides a rotating panel comprising approximately 8,000 households per wave, offering detailed information on household demographics, income, wealth, and financial assets. Our period of analysis spans from 2004 to 2020 and covers eight survey waves. We restrict our sample to mixed-gender households containing both male and female adults. Within this sample, 69% are male-headed households, and 31% are female-headed households.

3.2 Empirical Specification

The baseline empirical specification takes the following form:

$$y_{hrt} = \alpha + \beta FIS_{hrt} + \sum_{z=1}^{Z} \gamma_z Control_{z,hrt} + \delta_{rt} + \epsilon_t \tag{1}$$

where h, r, and t denote the household, NUTS-2 region of residence, and time (wave), respectively. The key independent variable is the **Female Income Share** (FIS), which serves as our measure of women's economic empowerment within the household. It is defined as the share of household income earned by women in the household relative to total household income:

$$FIS = \frac{\text{Total Female Income}}{\text{Total Household Income}} \tag{2}$$

The dependent variable, y_{hrt}, measures the household's allocation to risky assets, captured in two ways. First, we define the dummy variable *Risk Holding*, which takes the value of 1 if the household holds equities and/or private bonds, and 0 otherwise. Second, we restrict attention to households investing in financial assets and define FIS as the ratio of the value of equities and private bonds held in the portfolio to the total value of the household's financial assets.

The vector $Control_z$ includes a comprehensive set of household and individual-level variables: household income and its square, wealth quartiles, number of adults, number of children, average risk aversion, average age of adult household members and its square, average education of adult household members, marital status (married, divorced), homeownership, and employment, self-employment and retired status (The complete list of variables used in the empirical analysis, together with their definitions, descriptive statistics, and mean differences between male- and female-headed households, is reported in Tables A1 and A2 in the Appendix). We also include region×year fixed effects (δ_{rt}) to absorb local macroeconomic conditions and time-varying social norms, in the spirit of Guiso and Zaccaria (2023). We estimate pooled linear probability models and Tobit models, with standard errors clustered at the household level.

3.3 Baseline Results

Table 3 presents the baseline estimation results on the full sample of mixed-gender households. The coefficient on FIS is negative and highly significant in both risk-holding and risk-share specifications. Specifically, an increase in the female income share is associated with a 5.2-5.4 percentage-point decrease in the probability of holding risky assets and a substantial decrease in the overall portfolio share allocated to risky assets.

Table 3: Baseline results: FIS and asset Allocation

	Risk Holding		*Risk Share*	
FIS	-0.052***	-0.054***	-0.218***	-0.222***
	[0.006]	[0.007]	[0.0291]	[0.030]
Controls	inc/wealth	all	inc/wealth	all
Region×Year FEs	yes	yes	yes	yes
Observations	41,916	41,916	34,732	34,732
R^2/Pseudo R^2	0.210	0.225	0.197	0.216

Notes. Controls include: household income; household income squared; dummies for household wealth quartiles; average age of adults; average age of adults squared, average education of adults, HH employed, HH self-employed, HH retired, Partner employed, Partner self-employed; average risk aversion of adults, married status, divorced status, number of adults, number of children, homeownership. *** p<0.01.

4 Household Decision-Making: Theory and Empirical Evidence

This section studies how women's relative contribution to household income enters household financial decisions. We contrast two canonical views of intra-household decision-making. The first is a bargaining model, in the spirit of McElroy and Horney (1981), in which household choices aggregate the preferences of different household members, weighted by their relative bargaining power. The second is a unitary or dictatorship model, following the tradition of Becker (1974, 1981b), in which a single household member effectively controls financial decisions. The purpose of the exercise is not to provide a complete structural model of household behavior, but rather to derive sharp empirical implications that allow us to distinguish between these two views in the context of risky portfolio choice.

Consider a two-adult, mixed-gender household endowed with one unit of financial wealth. The household can allocate wealth between a risk-free asset with zero return and a risky asset s, whose return is distributed as $R_s \sim N(\mu_s, \sigma_s^2)$. Let α denote the share of wealth invested in the risky asset. Each adult household member $j \in \{F, M\}$, where F denotes the female member and M the male member, is characterized by a mean-variance utility function:

$$U_j(\alpha) = \alpha\mu_s - \frac{\gamma_j}{2}\alpha^2\sigma_s^2 - C \cdot \mathbb{1}_{\alpha>0}, \tag{3}$$

where γ_j captures individual risk aversion. Participation in the risky asset market also entails a fixed non-monetary cost C, so that investing in the risky asset requires paying $C \cdot \mathbb{1}_{\alpha>0}$.

4.1 The bargaining approach

Under the bargaining model, household utility is a weighted average of individual utilities:

$$V_H(\alpha) = \omega(FIS)U_F(\alpha) + [1 - \omega(FIS)]\, U_M(\alpha) - C \cdot \mathbb{1}_{\alpha>0}, \tag{4}$$

where FIS denotes the female income share. The bargaining weight of the female member increases with her income contribution, and, for expositional simplicity, we set $\omega(FIS) = FIS$. Substituting individual utilities into the household objective function gives:

$$V_H(\alpha) = \alpha\mu_s - \frac{\Gamma_H}{2}\alpha^2\sigma_s^2 - C \cdot \mathbb{1}_{\alpha>0}, \tag{5}$$

where household risk aversion is:

$$\Gamma_H = \gamma_M + FIS(\gamma_F - \gamma_M). \tag{6}$$

The optimal risky asset share is therefore:

$$\alpha^* = \begin{cases} 0 & \text{if } \Gamma_H > \widetilde{\Gamma} = \dfrac{\mu_s^2}{2C\sigma_s^2}, \\[2ex] \dfrac{1}{\Gamma_H}\dfrac{\mu_s}{\sigma_s^2} & \text{if } \Gamma_H \leq \widetilde{\Gamma} = \dfrac{\mu_s^2}{2C\sigma_s^2}. \end{cases} \tag{7}$$

Differentiating (7) with respect to FI, it is straightforward to verify that

$$\frac{\partial Prob(\alpha_H^* > 0)}{\partial FIS} \lesseqgtr 0 \;\text{ and }\; \frac{\partial \alpha_H^*}{\partial FIS} \lesseqgtr 0 \iff \gamma_F \gtreqless \gamma_M \tag{8}$$

Therefore, this framework delivers a simple testable implication. If household financial decisions result from a bargaining process in which members seek to have their preferences prevail, then when the female household member is more risk-averse than the male member, an increase in FIS raises household risk aversion and reduces both the probability of holding risky assets and the share of risky assets. If the female member is less risk-averse than the male member, the opposite should occur. If the two members have the same level of risk aversion, FIS should have no effect on household portfolio choices. Importantly, under this interpretation, the sign of the effect of FIS should depend on the relative risk aversion of women and men within the household, not on the gender of the person formally or informally in charge of financial decisions.

4.1.1 Measuring risk preferences

Testing this prediction requires information on the risk preferences of all adult household members. The SHIW provides information on risk attitudes only for the survey respondent, who is typically the household head and the individual responsible for financial decisions. In particular, risk aversion is elicited by a survey question that asks respondents to rank their preferences among alternative investment profiles, ranging from very high returns with a high risk of capital loss to low returns with no risk of loss. Responses are recorded on a four-point scale from 1 (high risk tolerance) to 4 (strong risk aversion).[1] By survey design, this question was administered only to the household head, leaving risk preferences systematically unobserved for all other adult members, including spouses and partners whose attitudes are central to our intra-household bargaining framework. Excluding these observations would preclude construction of the gender-differentiated, household-level risk measures our analysis requires.

We address this through MissForest (Stekhoven and Bühlmann, 2012), a machine learning imputation algorithm built on Random Forests (Breiman, 2001). Unlike mean substitution, which compresses variance and biases estimates toward zero (Little and Rubin, 2002)—or Multiple Imputation by Chained Equations, which requires a separately specified parametric model for each incomplete variable and risks misspecification (van Buuren and Groothuis-Oudshoorn, 2011), MissForest imposes no distributional assumptions, handles continuous, binary, and categorical variables within a single unified framework, and learns non-linearities and interactions directly from the data. These properties make it particularly well-suited to our setting, where the determinants of risk tolerance likely interact in complex ways across gender, age, and socioeconomic background.

4.1.2 The MissForest algorithm

The algorithm works as follows. A Random Forest builds a large ensemble of decision trees, each trained on a bootstrap sample of observations with a random subset of predictors

[1]Specifically, survey question C31 asks: "In managing your financial investments, would you say you have a preference for investments that offer: (a) very high returns, but with a high risk of losing part of the capital (1); (b) a good return, but also a fair degree of protection for the invested capital (2); (c) a fair return, with a good degree of protection for the invested capital (3); (d) low returns, with no risk of losing the invested capital (4)?"

considered at every split. For *risfin*, which takes $K = 4$ ordered values, the ensemble prediction is determined by majority vote across all B trees:

$$\hat{f}(\mathbf{x}^*) = \underset{k\in\{1,\ldots,K\}}{\arg\max} \sum_{b=1}^{B} \mathbf{1}\Big\{\hat{f}_b(\mathbf{x}^*) = k\Big\}. \tag{9}$$

Node splits maximise the reduction in Gini impurity, $\text{Gini}(N) = 1 - \sum_{k=1}^{K} \hat{p}_k^2$, where $\hat{p}_k$ is the proportion of observations in node N belonging to class k. Aggregating across many decorrelated trees substantially reduces prediction variance, and the out-of-bag (OOB) observations—those excluded from each tree's bootstrap sample—provide an unbiased estimate of predictive accuracy without requiring a separate validation set.

MissForest applies this framework iteratively. Missing *risfin* entries are first initialized with the modal category observed among household heads. The algorithm then trains a classification forest on the observed portion of the data, predicts values for the missing entries, and replaces the placeholders with updated estimates. This process repeats, with each cycle using the most recently imputed values for all variables, until the proportion of category labels that change between successive iterations,

$$\Delta_{\text{cat}}^{(t)} = \frac{1}{|M_{\text{cat}}|} \sum_{(i,j)\,\in\, M_{\text{cat}}} \mathbf{1}\Big\{\hat{X}_{ij}^{(t)} \neq \hat{X}_{ij}^{(t-1)}\Big\}\,, \tag{10}$$

falls below $\tau = 10^{-5}$, subject to a maximum of ten iterations.

The set of variables used for the imputation includes age, gender, education, individual income, marital status, household role, occupational status, employment sector, and two binary indicators: *fin work*, which identifies whether the household member is employed in the financial services sector (defined as occupations coded as *aspett* $\in \{6, 7\}$, according to the SHIW classification), and fin_study (a finance-related educational qualification, *tipolau* $\in \{1, 4, 5, 6\}$), which are strong predictors of financial risk attitudes (Christiansen et al., 2008; Calvet and Sodini, 2014). Before estimation, we correct an important SHIW data convention: the survey replicates the household head's *risfin* value across all household members by default. We reassign *risfin* to missing for all non-head individuals so that the algorithm trains exclusively on genuine survey responses. The procedure is implemented with 100 trees and a fixed random seed (`seed`$=$ 123) for full replicability, and is applied separately to each of the seven survey waves spanning 2006 to 2020, ensuring that imputed values reflect the distributional properties and covariate structure prevailing in each period rather than a pooled average that could obscure meaningful variation over time.

We assess the reliability of the imputed values using two complementary diagnostics applied wave by wave. First, we report the OOB classification accuracy produced by MissForest, an unbiased internal estimate of predictive performance that requires no separate holdout set. Second, we compute two-sample Kolmogorov-Smirnov (KS) statistics comparing the empirical distribution of observed *risfin* values (household heads with directly elicited responses) against the full distribution of imputed values across all adult members, as a practical measure of distributional distortion introduced by the imputation. Results are presented in Table 4, and the corresponding kernel density overlays for each wave are shown in Figure 1.

Table 4: MissForest imputation diagnostics by survey wave

Wave	OOB accuracy	KS statistic	Assessment
2006	0.650	0.142	Good
2008	0.671	0.170	Good
2010	0.563	0.117	Acceptable
2012	0.733	0.108	Strong
2014	0.699	0.196	Good
2016	0.656	0.151	Good
2020	0.583	0.114	Acceptable

Notes: OOB accuracy is $1 - \text{PFC}$, where PFC is the proportion of falsely classified observations from MissForest's internal out-of-bag estimation; the random-guessing benchmark for a four-category variable is 0.25. The KS statistic is the two-sample Kolmogorov-Smirnov statistic comparing the empirical distribution of observed values for household heads and directly elicited responses with the distribution of imputed values across all adult household members. *Strong* ≥ 0.70; *Good* 0.62–0.69; *Acceptable* 0.55–0.61.

OOB accuracy ranges from 0.563 in 2010 to 0.733 in 2012, comfortably exceeding the 0.25 benchmark of random classification across four categories in every wave. Five of the seven waves achieve accuracy of at least 0.650, confirming that the observed individual and household characteristics carry substantial predictive content for financial risk propensity. The two lower-performing waves, 2010 and 2020, coincide with periods of exceptional macroeconomic disruption, namely the aftermath of the global financial crisis and the COVID-19 pandemic, respectively, when the relationship between socioeconomic characteristics and risk preferences may have been temporarily less stable. Even in these waves, however, accuracy remains more than twice the random-guessing baseline. KS statistics range from 0.108 in 2012 to 0.196 in 2014, indicating that the cumulative distributions of observed and imputed *risfin* values differ by at most 20 percentage points across all waves. The associated p-values are all zero, a pattern that reflects the well-documented tendency of the KS test to flag even negligible differences as statistically significant with large samples, rather than as evidence of meaningful distributional distortion.

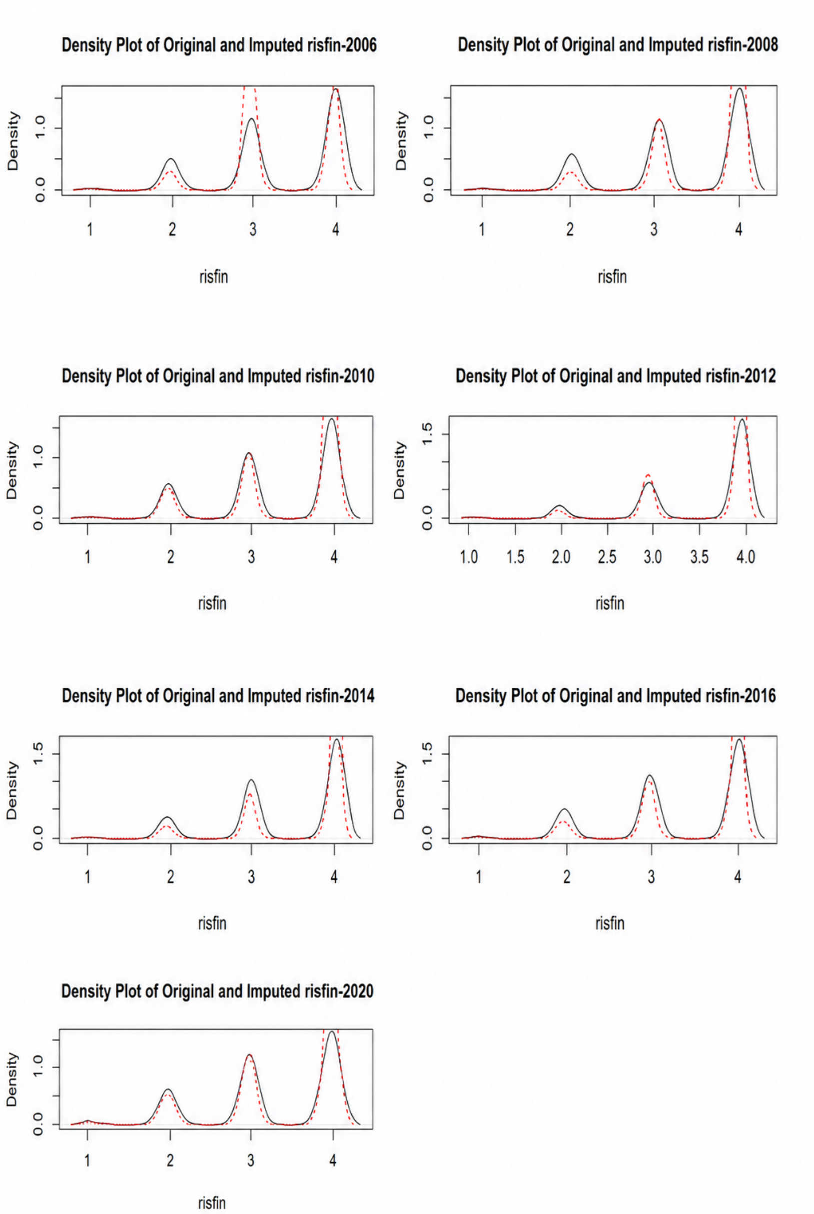


Figure 1: Kernel density plots of observed and imputed *risfin* values by survey wave

Notes. Each panel plots the kernel density of observed *risfin* values (solid black line, household heads with directly elicited responses) against the kernel density of imputed *risfin* values (dashed red line, all adult household members including non-heads) for the indicated survey wave. The four-point *risfin* scale runs from 1 (high risk tolerance) to 4 (strong risk aversion). The close correspondence between the two density curves in each wave indicates that the MissForest imputation preserves the marginal distribution of financial risk propensity without introducing systematic bias toward any single risk category.

The kernel density plots in Figure 1 provide more informative visual evidence: across all seven waves, the imputed density (dashed red line) closely follows the observed density (solid black line), reproducing both the bimodal structure of the distribution-, concentrated at categories 3 and 4, consistent with the broad risk aversion documented in the household finance literature, and the relative magnitudes of the two peaks. No wave exhibits systematic compression toward a single category or meaningful displacement of distributional mass, suggesting the imputation preserves the shape of the underlying risk preference distribution with good fidelity. Taken together, the OOB accuracy figures and the density comparisons support the reliability of the imputed *risfin* values across the full sample period.

Following imputation, we construct the partner's risk propensity, the gender-disaggregated average risk-preference for adult male and female members, and the overall household-level risk-preference measures used in the regression analysis. This allows us to incorporate the risk preferences of all household members into our analysis and aggregate them across genders, rather than relying exclusively on risk preferences reported by the household head. It also enables us to assess whether household financial investment decisions are consistent with a bargaining process among members with heterogeneous risk attitudes (McElroy and Horney, 1981; Browning and Chiappori, 1998; Mazzocco, 2004), or with a unitary framework in which within-household differences in risk preferences influence the choices of a benevolent household head (Becker, 1981a).

4.1.3 Evidence

To test whether the bargaining model can explain the observed pattern of household financial decisions, we classify households, using both observed and imputed measures of risk aversion, into three categories: adult women are, on average, more risk-averse than adult men, less risk-averse, or equally risk-averse. Formally, let $\Delta_{FM} = Avg(\gamma_F) - Avg(\gamma_M)$ denote the difference between the average risk aversion of female and male household members. Therefore, we estimate the following model:

$$\begin{aligned} y_{hrt} = \alpha + \beta_1 FIS \times \mathbb{1}_{\Delta_{FM}>0} + \beta_2 FIS \times \mathbb{1}_{\Delta_{FM}<0} + \\ \beta_3 FIS \times \mathbb{1}_{\Delta_{FM}=0} + \sum_{z=1}^{Z} \gamma_z Control_{z,hrt} + \delta_{rt} + \epsilon_t \end{aligned} \tag{11}$$

where $\mathbb{1}\Delta FM \lesseqgtr 0$ denotes three indicator variables that take the value of one when Δ_{FM} is less than, greater than, or equal to zero, respectively. The comparative statics in (8) imply that, under the bargaining model, $\beta_1 < 0$, $\beta_2 > 0$, and $\beta_3 = 0$, corresponding to cases in which $\Delta_{FM} > 0$, $\Delta_{FM} < 0$, and $\Delta_{FM} = 0$, respectively.

Table 5 reports the corresponding estimates. The evidence does not support the bargaining prediction. The coefficient of FIS is negative when women are more risk-averse than men, and it remains negative, albeit smaller in magnitude and less statistically significant, even when women are less risk-averse. More importantly, it remains negative and highly significant even when adult women and men within the household have the same level of risk aversion. In this latter case, when the preferences of household members coincide, the bargaining model predicts no effect of FIS on either risky-asset participation or portfolio shares. The persistence of a sizeable effect in this group is therefore difficult to reconcile with a model in which women's income share matters only through changes in the relative weight attached to female and male risk preferences.

Table 5: Bargaining model: Female income share and relative risk aversion

	Risk Holding	*Risk Share*
$FIS \times \mathbb{1}_{\Delta_{FM}>0}$	-0.099^{***}	-0.342^{***}
	(0.010)	(0.042)
$FIS \times \mathbb{1}_{\Delta_{FM}<0}$	-0.018^{*}	-0.056
	(0.011)	(0.045)
$FIS \times \mathbb{1}_{\Delta_{FM}=0}$	-0.043^{***}	-0.200^{***}
	(0.008)	(0.037)
Controls	Yes	Yes
Region×year fixed effects	Yes	Yes
Observations	41,916	34,732
R^2/Pseudo-R^2	0.227	0.218

Notes: The table reports pooled linear probability estimates for *Risk Holding* and pooled Tobit estimates for *Risk Share*. $\Delta_{FM} = Avg(\gamma_F) - Avg(\gamma_M)$ denotes the average female–male risk-aversion differential within the household. $\mathbb{1}_{\Delta_{FM}>0}$ identifies households in which adult women are, on average, more risk averse than adult men; $\mathbb{1}_{\Delta_{FM}<0}$ identifies households in which adult women are less risk averse; $\mathbb{1}_{\Delta_{FM}=0}$ identifies households in which adult women and men have the same average risk aversion. Standard errors are reported in parentheses. Controls are the same as in Table 3. In addition, we control for the average risk aversion of adult females and average risk aversion of adult males. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.2 The unitary household approach

We next consider a unitary model of household decision-making. In this framework, portfolio choices are not the outcome of a weighted average of individual preferences but rather reflect the preferences, beliefs, and confidence of the household member (the head of the household) who effectively manages financial investments. This view is consistent with a Beckerian representation of the household as if it maximized a single objective function, possibly that of an altruistic, benevolent household head (Becker, 1981b,a). In the present context, the key implication is that the effect of FIS should depend on who controls financial decisions. If the household head is female, a higher female income share may reinforce the decision-maker's confidence and authority in financial matters or influence her level of altruism. If the household head is male, an increase in the female income share (FIS), which corresponds to a decline in the male income share 1 -FIS, may instead weaken the perceived dominance of the male decision-maker, reduce his confidence in portfolio choices, or alter his willingness to accommodate the preferences of other household members. Regardless of the specific channel through which FIS operates, the key implication is that, within a unitary household framework, the effects on investment choices should exhibit opposite signs depending on whether the household head is male or female.

We formalize this idea by allowing the household objective to differ according to the gender of the financial decision-maker:

$$V_{FH} = f_F(U_1, \ldots, U_n, FIS) labeleq : VFH \tag{12}$$

$$V_{MH} = f_M(U_1, \ldots, U_n, 1 - FIS) \tag{13}$$

where V_{FH} and V_{MH} denote the objective functions of female-headed and male-headed households, respectively. Unlike in the bargaining model, the sign of the effect of FIS is not pinned down by the relative risk aversion of women and men. Instead, it is expected to vary with the identity of the household head. However, a central prediction of the unitary model is that the effect of women's income share on risky investment should be different, and potentially opposite, in male- and female-headed households

We estimate Heckman selection models for household headship and portfolio outcomes, using widow or widower status as an exclusion restriction. Consistent with a unitary framework, the results reported in Table 6 show a sharp asymmetry. First, the probability that the household head is a woman increases with women's contribution to total household income (for brevity, the selection equation estimates are not reported in the table). Turning to the outcome equations, in male-headed households, FIS has a negative and statistically significant effect on both risky asset holding and the share of risky assets. In female-headed households, by contrast, the effect is positive and statistically significant. Quantitatively, an increase in the female income share reduces the probability of holding risky assets by 11 percentage points in male-headed households and increases it by 8.1 percentage points in female-headed households. The same sign reversal is observed for the share of risky assets.

Table 6: Unitary model: Female income share by gender of household head

	Male headship		*Female headship*	
	Risk Holding	*Risk Share*	*Risk Holding*	*Risk Share*
FIS	-0.110^{***}	-0.085^{***}	0.081^{***}	0.058^{***}
	(0.020)	(0.016)	(0.023)	(0.022)
λ	0.008	0.017^{*}	0.025^{*}	0.003
	(0.013)	(0.009)	(0.014)	(0.012)
Controls	Yes	Yes	Yes	Yes
Region×year fixed effects	Yes	Yes	Yes	Yes
Observations	28,876	24,335	13,040	10,397

Notes: The table reports Heckman selection estimates for risky asset holding and risky asset shares by gender of the household head. Male- and female-headed households are households in which the person in charge of financial decisions is male or female, respectively. Controls are the same as in Table 5. Widow/widower status is used as an exclusion restriction. Standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The differential impact of women's economic empowerment on household investment choices is difficult to reconcile with a pure intra-household bargaining mechanism, but consistent with a unitary model in which the identity of the financial decision-maker plays a central role. The evidence suggests that women's contribution to household income matters not only through the weight their preferences carry in shaping household investment decisions, but also through the effect on the allocation of financial decision rights. The evidence suggests that women's contribution to household income matters not only through the weight their preferences carry in shaping household investment decisions, but also through its effect on the allocation of authority in financial decision-making. In other words, the female income share influences portfolio choices through the household decision-making structure: the same change in income composition has opposite implications depending on whether financial management is effectively entrusted to a man or to a woman.

5 Mechanisms

In this section, we examine the mechanisms through which the female income share may affect household investment decisions in a unitary household decision-making framework. We focus on two channels. The first is an *altruism channel*. Under this mechanism, the female income share affects portfolio choices by altering the extent to which the household head internalizes the utility of the other adult household member. The second is a *self-confidence channel*. Under this mechanism, a greater female contribution to household income may affect the head of household's self-confidence. A larger income contribution may increase the perceived authority and confidence of the household head when the head is a woman, while it may weaken the confidence or relative authority of a male head. These two channels generate distinct empirical implications, allowing us to distinguish between them.

5.1 The altruism channel

5.1.1 Theory

The altruism mechanism builds on a unitary representation of household financial decisions, like in equations (**??**) and (13), in which the household head chooses the portfolio but may assign positive weight to the partner's utility. In particular, assuming that household members are characterized by mean-variance utility functions, in female-headed households, the female head maximizes

$$\begin{aligned} V_{FH} &= U_F + a_F(FIS)U_M - C \cdot \mathbb{1}_{\alpha>0} \\ &= [1+\alpha(FIS)]\mu_s + \frac{\gamma_F + a_F(FIS)\gamma_M}{2}\alpha^2\sigma_s^2 - C \cdot \mathbb{1}_{\alpha>0} \end{aligned} \tag{14}$$

while in male-headed households, the male head maximizes

$$\begin{aligned} V_{MH} &= U_M + a_M(1-FIS)U_F - C \cdot \mathbb{1}_{\alpha>0} \\ &= [1+\alpha(1-FIS)]\mu_s + \frac{\gamma_M + a_M(1-FIS)\gamma_F}{2}\alpha^2\sigma_s^2 - C \cdot \mathbb{1}_{\alpha>0}, \end{aligned} \tag{15}$$

where $a_F(FIS)$ and $a_F(1-FIS)$ capture the degree of altruism of the female and male head, respectively. The optimal risky asset allocation in female- and male-headed households are

$$\alpha_{FH}^{*} = \begin{cases} 0 & \text{if } \Gamma_{FH} > \widetilde{\Gamma}_{FH} = \frac{[(1+a_F(FIS))\mu_s]^2}{2C\sigma_s^2} \\ \frac{1+a_F(FIS)}{\Gamma_{MH}}\frac{\mu_s}{\sigma_s^2} & \text{if } \Gamma_{FH} \leq \widetilde{\Gamma}_{FH} = \frac{[(1+a_F(FIS))\mu_s]^2}{2C\sigma_s^2} \end{cases} \tag{16}$$

$$\alpha_{MH}^{*} = \begin{cases} 0 & \text{if } \Gamma_{MH} > \widetilde{\Gamma}_{MH} = \frac{[(1+a_M(1-FIS))\mu_s]^2}{2C\sigma_s^2} \\ \frac{1+a_M(1-FIS)}{\Gamma_{MH}}\frac{\mu_s}{\sigma_s^2} & \text{if } \Gamma_{FH} \leq \widetilde{\Gamma}_{MH} = \frac{[(1+a_M(1-FIS))\mu_s]^2}{2C\sigma_s^2} \end{cases} \tag{17}$$

where $\Gamma_{FH} = \gamma_F + a_F(FIS)\gamma_M$ and $\Gamma_{MH} = \gamma_M + a_M(1-FIS)\gamma_F$ denote the effective risk-aversion parameters. Differentiating (16) and (17) with respect to FIS, it is straightforward to verify that the effect of the share of income produced by females on risky asset holding and risky asset shares depends on the interaction between altruism and the gender gap in risk aversion:

$$sign\frac{\partial Prob(\alpha_{FH} > 0)}{\partial FIS} = sign\frac{\partial \alpha_{FH}^{*}}{\partial FIS} = sign\frac{\partial a_F}{\partial FIS}(\gamma_F - \gamma_M) \tag{18}$$

$$sign\frac{\partial Prob(\alpha_{MH} > 0)}{\partial FIS} = sign\frac{\partial \alpha_{MH}^{*}}{\partial FIS} = -sign\frac{\partial a_M}{\partial FIS}(\gamma_F - \gamma_M) \tag{19}$$

Although the female income share (FIS) may exert an ambiguous effect on household asset allocation, depending on whether altruism increases or decreases with it, the altruism channel delivers a clear, testable implication: if adult females and males are equally risk averse, FIS should have no effect on the asset allocation of either female- or male-headed households. Accordingly, if the female income share affects household portfolios primarily through the formation of altruistic preferences, its effect should vanish in the subsample of households in which adult women and men exhibit the same average level of risk aversion.

5.1.2 Evidence

To assess the relevance of the altruism channel in explaining the opposite effects of FIS on the financial investment choices of male- and female-headed households, we estimate a Heckman selection model on the restricted sample of households in which adult women and men exhibit the same average level of risk aversion. In this setting, the coefficient on FIS

should be statistically indistinguishable from zero for both risky-asset holdings and risky-asset shares. Table 7 reports the results. In male-headed households, the coefficient on FIS remains negative and statistically significant. An increase in the female income share reduces both the likelihood of holding risky assets and the share of risky assets, even when adult women and men exhibit the same average level of risk aversion. Similarly, in female-headed households, the coefficient on FIS remains positive and statistically significant for both outcomes.

Table 7: Altruism mechanism: households with equal average risk aversion

	Male headship		*Female headship*	
	Risk holding	*Risk share*	*Risk holding*	*Risk share*
FIS	-0.103^{***}	-0.078^{***}	0.089^{***}	0.091^{***}
	(0.030)	(0.024)	(0.033)	(0.033)
λ	0.005	0.012	0.042^{**}	0.029^{*}
	(0.019)	(0.014)	(0.020)	(0.017)
Controls	Yes	Yes	Yes	Yes
Region × year FEs	Yes	Yes	Yes	Yes
Observations	11,702	9,630	5,418	4,241

Notes. Heckman selection models estimated on the subsample of households in which adult women and adult men have the same average risk aversion. Control and exclusion restriction are the same as in Table 6. Robust standard errors in brackets. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The persistence of significant effects in this restricted sample is inconsistent with the central prediction of the altruism channel. If the female income share operated only by changing the weight attached to the partner's utility, it should not affect risky investment when male and female adults have identical average risk preferences. Instead, the coefficient on FIS remains large, statistically significant, and opposite in sign across male- and female-headed households. This pattern suggests that the female income share affects portfolio choices through a mechanism that goes beyond reweighting male and female risk preferences within the household objective function.

5.2 The self-confidence channel

5.2.1 Theory

The second mechanism links the female income share to the household head's confidence in their own assessment of risky assets. Portfolio choices depend not only on "true" risk preferences, but also on the decision-maker's perceived ability to evaluate the distribution of risky returns. A more confident household head may perceive the investment environment as less uncertain, process financial information more effectively, or feel more entitled to take responsibility for risky financial decisions without taking into account preferences and opinions that differ from their own.

To parsimoniously capture this mechanism within our mean–variance framework, we assume that a lack of confidence in one's own authority and judgment amplifies the negative weight assigned to investment risk in the household head's utility function

$$V_{FH} = \alpha\mu_s - \frac{\gamma_F}{2}\frac{1}{c_F(\phi, FIS)}\alpha^2\sigma_s^2 \tag{20}$$

$$V_{MH} = \alpha\mu_s - \frac{\gamma_M}{2}\frac{1}{c_M(\phi, 1-FIS)}\alpha^2\sigma_s^2 \tag{21}$$

where $c_F(\phi, FIS)$ denotes the confidence of a female household head in evaluating risky asset returns, and $c_M(\phi, 1-FIS)$ the corresponding confidence of a male household head, both of which increase with their contribution to household income and their financial expertise, such as employment in the financial sector.

In this case, the optimal portfolio allocation for female- and male-headed M-Households is

$$\alpha^*_{FH} = \begin{cases} 0 & \text{if} \quad \gamma_F > \tilde{\gamma}_F = \frac{c_F(\phi,FIS)\mu_s^2}{2C\sigma_s^2} \\ \frac{c_F(\phi,FIS)}{\gamma_F}\frac{\mu_s}{\sigma_s^2} & \text{if} \quad \gamma_F \leq \tilde{\gamma}_F = \frac{c_F(\phi,FIS)\mu_s^2}{2C\sigma_s^2} \end{cases} \tag{22}$$

$$\alpha^*_{MH} = \begin{cases} 0 & \text{if} \quad \gamma_M > \tilde{\gamma}_M = \frac{c_M(\phi,1-FIS)\mu_s^2}{C\sigma_s^2} \\ \frac{c_M(\phi,1-FIS)}{\gamma_M}\frac{\mu_s}{\sigma_s^2} & \text{if} \quad \gamma_M \leq \tilde{\gamma}_M = \frac{c_M(\phi,1-FIS)\mu_s^2}{C\sigma_s^2} \end{cases} \tag{23}$$

From (22) and (23), and consistent with the evidence reported in Table 6, an increase in the female share of household income raises both the likelihood that a female household head invests in risky assets and the share of such assets held, while exerting a negative effect on the risky investment choices of a male household head. Moreover, depending on whether the contribution to household income acts as a substitute for or a complement to financial experience in shaping the household head's confidence, the impact of FIS may either amplify or attenuate the positive effect of the household head's level of financial sophistication on risk-taking. In either case, its effect should be opposite for male and female household heads.

5.2.2 Evidence

We test this prediction by augmenting the Heckman selection model with an interaction between FIS and an indicator variable, $HFIN$, which equals 1 if the household head is employed in the financial sector. Employment in finance is used as a proxy for financial sophistication and confidence in asset evaluation.

Table 8 reports the results. In male-headed households, the coefficient on FIS is negative and statistically significant, while its interaction with $HFIN$ is positive and statistically significant for risky asset holding and positive, though not statistically significant, for the risky asset share. In female-headed households, the pattern is reversed: the coefficient on FIS is positive and statistically significant, whereas the interaction between FIS and $HFIN$ is negative, particularly strong, and statistically significant for the risky asset share. These findings are consistent with a self-confidence mechanism underlying the household head's unitary investment decisions. Employment in the financial sector and a higher share of household income both increase the likelihood of investing in risky assets and the portfolio share allocated to such assets. Moreover, the opposite signs of the interaction between FIS and $HFIN$ across male- and female-headed households are consistent with the two acting as substitutes: financial sophistication reduces the marginal effect of income shares on the head of household's self-confidence. Accordingly, when the household head is already highly confident in their ability to accurately assess the risks of financial assets, changes in relative income contributions exert a more limited effect on portfolio choices.

6 Conclusions

This paper examines whether and how women's contributions to household income affect household financial investment decisions. The evidence shows that the female income share is an important determinant of household portfolio allocation. On average, a higher female income share is associated with lower exposure to risky assets. However, this average effect masks substantial heterogeneity by the gender of the household head responsible for financial decisions.

The main results indicate that the effect of the female income share does not vary systematically with the relative risk aversion of women and men within the household, which is inconsistent with the predictions of a bargaining model. Instead, it varies sharply with the

Table 8: Self-confidence mechanism: interaction with financial-sector employment

	Male headship		Female headship	
	Risk holding	Risk share	Risk holding	Risk share
FIS	-0.110^{***}	-0.102^{***}	0.099^{***}	0.053^{**}
	[0.027]	[0.020]	[0.026]	[0.026]
$HFIN$	0.156^{***}	0.099^{***}	0.305^{***}	0.228^{***}
	[0.021]	[0.021]	[0.073]	[0.055]
FIS $\times$ $HFIN$	0.140^{*}	0.037	-0.168	-0.214^{**}
	[0.078]	[0.055]	[0.119]	[0.090]
λ	0.027	0.035^{***}	0.047^{***}	0.018
	[0.019]	[0.012]	[0.016]	[0.013]
Controls	Yes	Yes	Yes	Yes
Region $\times$ year FEs	Yes	Yes	Yes	Yes
Observations	28,876	24,335	13,040	10,397

Notes. Heckman selection models augmented with the interaction between the female income share and $HFIN$, an indicator equal to one when the household head is employed in the financial sector. Control and exclusion restriction are the same as in Table 6. Robust standard errors in brackets. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

gender of the financial decision-maker. In male-headed households, a higher female income share reduces both the probability of holding risky assets and the share of risky assets in the household portfolio. In female-headed households, the same increase has a positive effect on both outcomes. This pattern is consistent with a unitary model of household financial behavior in which the identity of the person who holds the purse strings is central.

The mechanism analysis further shows that the effect of the female income share is unlikely to operate through altruism. When the sample is restricted to households in which adult women and men have the same average risk aversion, the effect of the female income share remains statistically significant and has the opposite sign across male- and female-headed households. By contrast, the evidence supports a self-confidence channel. The effect of the female income share is attenuated when the household head is employed in the financial sector, suggesting that income shares matter less when the decision-maker already possesses financial expertise and confidence in asset evaluation.

Overall, women's economic empowerment affects household financial decisions not only by changing the distribution of resources within the household but also by interacting with the allocation of financial decision rights. The same increase in women's contribution to household income can lead to more or less financial risk-taking depending on who manages the household portfolio.

References

Addoum, J. M. (2017). Household portfolio choice and retirement. *Review of Finance*, 21(4):1641–1683.

Addoum, J. M., Kurnaz, M., and Veldkamp, L. (2016). Household portfolio diversification. Working Paper 22416, National Bureau of Economic Research.

Arano, K., Parker, C., and Terry, R. (2010). Gender differences in financial risk tolerance. *Financial Services Review*, 19(2):141–155.

Barber, B. M. and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116(1):261–292.

Becker, G. S. (1974). A theory of marriage: Part ii. *Journal of Political Economy*, 82(2, Part 2):S11–S26.

Becker, G. S. (1981a). Altruism in the family and selfishness in the market place. *Economica*, 48(1):1–15.

Becker, G. S. (1981b). *A Treatise on the Family*. Harvard University Press.

Bertocchi, G., Brunetti, M., and Torricelli, C. (2011). Marriage and stock market participation: Who call the shots? *Journal of Economic Behavior & Organization*, 80(3):562–578.

Bertocchi, G., Brunetti, M., and Torricelli, C. (2014). Who holds the purse strings within the household? the determinants of intra-household financial decision-making. *Review of Economics of the Household*, 12:651–678.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Browning, M. and Chiappori, P.-A. (1998). Efficient intra-household allocations: A general characterisation and empirical tests. *Econometrica*, 66(6):1241–1278.

Calvet, L. E. and Sodini, P. (2014). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *Journal of Finance*, 69(2):867–906.

Charness, G. and Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1):50–58.

Christiansen, C., Joensen, J. S., and Rangvid, J. (2008). Are economists more likely to hold stocks? *Review of Finance*, 12(3):465–496.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.

Dominitz, J. and Manski, C. F. (2007). Expected equity returns and portfolio decisions of older americans. *Journal of the European Economic Association*, 5(2-3):552–560.

Gomes, F., Haliassos, M., and Ramadorai, T. (2021). Household finance. *Journal of Economic Literature*, 59(3):919–1000.

Gu, R., Peng, C., and Zhang, W. (2024). The gender gap in household bargaining power: A revealed-preference approach. *The Review of Financial Studies*, page hhae039.

Guiso, L. and Zaccaria, T. (2023). Gender, social norms, and household portfolio allocation. Discussion paper, CEPR.

Jianakoplos, N. A. and Bernasek, A. (1998). Are women more risk averse? *Economic Inquiry*, 36(4):620–630.

Jianakoplos, N. A. and Bernasek, A. (2008). Family financial risk taking when the wife earns more. *Journal of Family and Economic Issues*, 29(2):289–306.

Ke, D. (2021). Who wears the pants? gender identity norms and household financial decisions. *Journal of Finance*, 76(3):1389–1425.

Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data.* Wiley, Hoboken, NJ, 2nd edition.

Love, D. A. (2010). The effects of marital status and children on savings and portfolio choice. *Review of Financial Studies*, 23(1):385–432.

Lusardi, A. and Mitchell, O. S. (2008). Planning and financial literacy: How do women fare? *American Economic Review*, 98(2):413–417.

Maura, G. (2022). Intra-household bargaining and household financial decisions. Technical report, Working Paper.

Mazzocco, M. (2004). Savings, risk sharing, and preferences for risk. *American Economic Review*, 94(4):1169–1182.

McElroy, M. B. and Horney, M. J. (1981). Nash-bargained household decisions: Toward a generalization of the theory of demand. *International Economic Review*, 22(2):333–349.

Stekhoven, D. J. and Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1):112–118.

Sunden, A. E. and Surette, B. J. (1998). Gender differences in the allocation of assets in retirement savings plans. *American Economic Review*, 88(2):207–211.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011). `mice`: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3):1–67.

Yilmazer, T. and Lich, S. (2015). Household portfolio choice and management. *Review of Economics of the Household*, 13(2):275–312.

Appendix: additional tables

Table A1: Variable definitions and statistics

Variable	Definition	N	Mean (SD)
Risk holding	Binary indicator = 1 if the household holds any risky financial assets (equities or private bonds), 0 otherwise.	61675	0.14 (0.35)
Risk share	Share of risky assets in total financial assets. Continuous variable in [0, 1].	49851	0.10 (0.24)
Core economic covariates			
FIS	Female income share in household: yindex2 = fem_y / (fem_y + mal_y), where fem_y is total income of all female members and mal_y is total income of all male members. Ranges from 0 to 1; 0.5 indicates income parity. Proxies the relative economic bargaining power of women within the household.	61513	0.41 (0.38)
logy	Natural logarithm of total household income: logy = log(y).	61508	10.19 (0.72)
logy2	Squared natural log of household income	61508	104.3 (14.3)
Labor market status			
employed_2	Binary indicator = 1 if the household head is a wage or salaried employee, 0 otherwise.	61675	0.32 (0.47)
selfemployed_2	Binary indicator = 1 if the household head is self-employed, 0 otherwise.	61675	0.10 (0.30)
retired_2	Binary indicator = 1 if the household head is retired, 0 otherwise.	61675	0.36 (0.48)
partner_employed_1	Binary indicator = 1 if the household head's partner is a wage or salaried employee, 0 otherwise: partner_employed_1 = 1 if partner_employ_status = 1.	61675	0.19 (0.40)
partner_selfemployed	Binary indicator = 1 if the household head's partner is self-employed, 0 otherwise: partner_selfemployed = 1 if partner_employ_status = 2.	61675	0.05 (0.21)
Education and financial literacy			
avg_impu_risfin_adult	Average risk propensity of all adult members	61675	3.44 (0.58)
avg_edu_all_adults	Average education of all adult members (age 18)	61675	3.77 (1.55)
Wealth controls			
quantile2wealth	Binary indicator = 1 if the household belongs to the second quartile of the net wealth distribution (w_quartile = 2), 0 otherwise.	61675	0.25 (0.43)
quantile3wealth	Binary indicator = 1 if the household belongs to the third wealth quartile (w_quartile = 3), 0 otherwise.	61675	0.25 (0.43)
quantile4wealth	Binary indicator = 1 if the household belongs to the fourth (top) wealth quartile (w_quartile = 4), 0 otherwise.	61675	0.25 (0.43)
Household structure			
female	Binary indicator = 1 if the household head is female, 0 otherwise.	61,675	0.42 (0.49)
num_adult	Number of adult household members	61675	2.09 (0.96)
num_kid	Number of children (age < 18) in household	61675	0.35 (0.73)
married	Binary indicator = 1 if the household head is legally married, 0 otherwise.	61675	0.61 (0.49)
divorced	Binary indicator = 1 if the household head is divorced, 0 otherwise.	61675	0.08 (0.27)
homeowner	Binary indicator = 1 household owns home, 0 otherwise.	61675	0.72 (0.45)
Life Cycle			
avg_age_all_adults	Average age (in years) of all adult members (age 18) in household	61675	55.72 (16.4)
avg_age_all_adults_2	Squared average age of all adults	61675	3374 (1913)

Notes. The sample consists of all household-year observations from the SHIW waves 2006–2020. Standard deviations are in parentheses. `risk_share` has a smaller N because it is defined only for households with positive total financial assets. Wealth quartile dummies sum to one; their means equal 0.25 by construction. The omitted reference categories are the first wealth quartile, not employed (labour market status), and single (marital status).

Table A2: Univariate analysis.

	Male	**Female**	**Diff.**	***t*-stat.**
Risk holding	0.18	0.10	0.08***	(28.36)
Risk share	0.11	0.07	0.04***	(20.06)
FIS	0.17	0.73	-0.56***	(-249.36)
logy	10.32	10.01	0.31***	(53.68)
logy2	106.95	100.65	6.30***	(55.57)
employed_2	0.36	0.27	0.09***	(23.74)
selfemployed_2	0.14	0.05	0.09***	(37.82)
retired_2	0.43	0.26	0.17***	(44.84)
partner_employed_1	0.21	0.17	0.04***	(13.09)
partner_selfemployed	0.04	0.05	-0.01***	(-7.24)
avg_impu_risfin_adult	3.38	3.50	-0.12***	(-25.05)
avg_edu_all_adults	3.93	3.56	0.37***	(29.27)
quantile2wealth	0.23	0.28	-0.05***	(-13.94)
quantile3wealth	0.25	0.24	0.01*	(2.47)
quantile4wealth	0.30	0.18	0.12***	(35.31)
num_adult	2.27	1.86	0.40***	(52.60)
num_kid	0.38	0.31	0.07***	(11.49)
married	0.77	0.39	0.37***	(98.94)
divorced	0.05	0.12	-0.06***	(-27.71)
homeowner	0.74	0.68	0.06***	(17.33)
avg_age_all_adults	54.24	57.74	-3.50***	(-25.79)
avg_age_all_adults_2	3180.73	3639.60	-458.87***	(-28.88)
Observations	35,647	26,028	61,675	

Notes: The sample comprises all household-year observations from the SHIW waves 2006–2020. The table reports means for male-headed and female-headed households separately, the difference between them, and the two-sample t-statistic (unequal variances) in parentheses. Significance levels: $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.