UNIVERSITAT DE BARCELONA

Essays on education, marriage and social classes

Hamza Benazzi

PhD in Economics

Essays on education, marriage, and social classes.

Hamza Benazzi

UNIVERSITAT DE BARCELONA

PhD in Economics

Thesis tittle:

Essays on education, marriage, and social classes.

PhD candidate:
Hamza Benazzi

Advisor:
Xavier Raurich

Date:
December 2024

UNIVERSITAT DE BARCELONA

To my beloved wife,
to whom I owe the leaping delight
that quickens my senses in our waking time
and the rhythm that governs the repose of our sleeping time,
the breathing in unison.

— Adapted from T.S. Eliot

I could not have finished this without you.

Acknowledgement

I am deeply grateful to my advisor, Xavier Raurich, for his invaluable guidance both before and during my PhD journey. One of the pivotal reasons I decided to pursue a PhD was the inspiration I found in his class during my master's studies. His eloquence and remarkable ability to simplify complex economic theories ignited my passion for the field, for which I will always be thankful. However, my gratitude goes beyond academics. During one of the most challenging periods of my life, following the loss of my sister, Xavier showed extraordinary empathy and support. His kindness and understanding during that time meant more to me than words can express.

I am profoundly grateful to Calin Arcalean and Ioana Shipou for coauthoring one of the chapters of this thesis with me. Collaborating with them on this project was an invaluable learning experience that significantly enhanced my research skills. I am a far better researcher now than I was when I first had the privilege of working with them. Their support extended beyond our collaboration, opening doors to opportunities that I firmly believe will serve as a springboard for my career. What stands out most is their willingness to help, driven not by obligation but by genuine generosity and kindness, for which I am deeply appreciative.

I would also like to express my heartfelt gratitude to Luiz Brotherhood for his invaluable support with my job market paper. His guidance was instrumental in advancing my PhD journey, and I am deeply appreciative of his contributions.

I would like to thank Jordi Roca for his tireless efforts in making the PhD process as seamless as possible. He is truly a cornerstone of the UB School of Economics PhD program, and his dedication does not go unnoticed.

I am profoundly grateful to the friends and colleagues I had the privilege of meeting along this journey. Each of them, in their own unique way, made the path more bearable and enriched the experience. I am especially thankful to Nabil, whose support and presence during the darkest days of my life provided me with comfort that I will always cherish.

When it comes to my parents, words will never suffice to express my gratitude. They sacrificed their youth, their money, and their comforts to ensure the best for me and my sisters. They endured loss and hardship, always putting our education and well-being above their own. I would not have become the man I am now without

them. For this, and their boundless love and selflessness, I am forever indebted.

During my PhD, I faced the devastating loss of my sister, a pain that could have easily derailed my journey. In those darkest moments, my wife, Kawtar, became my anchor and my greatest source of motivation. Her unwavering support, love, and encouragement gave me the strength to keep going. Without her, this work would never have seen the light of day.

Abstract

This thesis investigates the mechanisms through which economic inequality is created and perpetuated. Drawing from diverse economic theories and empirical evidence, it examines three interconnected topics that contribute to the persistence of inequality.

The first chapter explores the impact of government spending on education, revealing a paradox where increased public education investment can disproportionately disadvantage middle-income households by influencing their decisions to opt for private schooling. Using an overlapping generations model, it highlights the implications of such dynamics on intergenerational mobility and income disparities.

The second chapter analyzes the relationship between income inequality and local public education funding in U.S. school districts, demonstrating that higher inequality leads to reduced local funding, particularly in economically disadvantaged areas. An instrumental variable strategy is employed to address endogeneity concerns. The findings emphasize the need for equitable education policies that mitigate these adverse effects.

The final chapter investigates shifting marriage trends in the U.S., focusing on the role of wage structures and working hours in shaping household dynamics and inequality. It develops a general equilibrium model to analyze the interplay of economic and temporal factors in marital decisions. The findings indicate that changes in work hours play a crucial role in explaining recent marriage trends.

Overall, this dissertation contributes to understanding how education, redistribution, and marriage interact to sustain inequality and proposes avenues for future research and policy interventions aimed at mitigating inequality and fostering mobility.

Keywords: Education funding, Inequality, Intergenerational transfers, Marriage, Political Economy, Public and Private education, Social classes, Wage structure, Working hours.

Contents

List of Figures

List of Tables

1 Introduction

Introduction

Inequality has been and still is one of the biggest concerns for economists. It is the consequence of different mechanisms, including differences in access to education and structural disparities in wages and wealth distribution. Moreover, inequality can perpetuate and even exacerbate itself through feedback loops in mechanisms of redistribution. For example, wealthier individuals and groups can shape tax policies, social programs, and labor laws in ways that favor their interests through political influence, increasing inequality further. Additionally, unequal access to resources like quality education and job opportunities means that future generations inherit and often intensify these disparities. Thus, inequality doesn't merely persist; it can actively generate more inequality, creating a cycle that is difficult to break without intentional, systemic intervention.

There is ample evidence in the literature that economic, institutional, demographic, and cultural factors both cause and contribute to inequality. Human capital is among these mechanisms. Becker (1994) and Mincer (1974) highlight the central role of human capital accumulation, arguing that differences in education and skill acquisition underpin labor market disparities. Hanushek and Woessmann (2011) show how variations in human capital accumulation across countries are directly linked to inequality and economic performance. Goldin and Katz (2008) argue that inequality in the U.S. labor market has been driven by a "race" between the demand for skills caused by technological advancements (e.g., mechanization, automation, and IT innovations) and the supply of skills streamlined by the expansion of educational attainment, particularly high school and college education Wage inequality is determined by the balance between these two forces. Wage gaps remain small when the supply of skilled workers keeps up with technological advancements. However, wage disparities increase when technological advancements surpass the growth of education.

At the institutional level, policies governing taxation, social spending, and labor market regulations significantly influence the distribution of income and wealth, having a direct impact on inequality. According to Piketty and Goldhammer (2014), Piketty and Goldhammer (2020) and Piketty and Saez (2013), inequality in advanced economies has grown due to decreasing tax progressivity and increasing dominance of capital income over labor income. Furthermore, the weakening of union power (Freeman, 1992) and the prevalence of precarious work arrangements due to the rise of superstar firms (Autor et al., 2020) reduced wage growth for low and middle-income workers, resulting in polarization. Additionally, globalization in advanced economies has contributed to the offshoring of low-skill employment, further polarizing labor markets and expanding income inequality (Helpman et al., 2010).

Cultural and demographic factors also play an important role in shaping inequality. Solon (1992) and Chetty et al. (2014) show the persistence of inequality through intergenerational physical and nonphysical transfers such as bequests and investment in education. Greenwood et al. (2014) document an increase in assortative mating, which increases inequality by concentrating wealth within high-income households. de la Croix and Doepke (2003) argue that differential fertility between poor and rich reinforces income inequality across generations as lower-income families tend to have higher fertility rates and invest less in the education of their kids, while higher-income families have lower fertility and invest more in education.

Feedback loops inherent in inequality intensify its persistence. Wealthier individuals and groups often use their economic power to influence political and institutional outcomes (Stiglitz, 2012). This influence can shape tax policies, social programs, and labor laws in ways that protect and extend their wealth. Furthermore, unequal access to key services, such as quality education, slows upward mobility and increases inequality. Many papers in the literature, including (Jerrim and Macmillan, 2015), show that countries with higher income inequality tend to have lower levels of intergenerational mobility because of the education mechanism, which eventually leads to higher inequality.

This rich body of literature serves as the foundation for the themes explored in this dissertation, which investigates the mechanisms through which inequality is perpetuated. Each chapter sheds light on specific dimensions of inequality, from education and redistribution to labor changes and marriage. In this dissertation, I explore three distinct yet interconnected mechanisms affecting inequality. First, I examine how government intervention can sometimes inadvertently worsen inequality. Second, I investigate how inequality can perpetuate and deepen itself through mechanisms of redistribution. Finally, I analyze how social preferences, such as the desire for traditional bonds like marriage, can be shaped and influenced by different economic factors, which eventually might contribute to and reinforce economic inequality.

In Chapter 2, I use repeated cross-sections from IPUMS USA data spanning the years 2000 to 2019 and employ a logit discrete choice model as my empirical methodology. I document a heterogeneous response of parents towards sending their children to private schools, finding that middle-class parents are more elastic to changes in government spending on education compared to other income groups. Motivated by this finding, I investigate the impact of public versus private education choices on income inequality, focusing on the middle class within a theoretical framework. Using an overlapping generations model, we examine how parental decisions on education investment (public versus private) affect their children's future income and thus perpetuate or reduce inequality. We specifically consider three government policies: increasing public education spending per capita, raising income tax rates, and improving the quality of public education. Our model reveals that increased government spending in public education, paradoxically, reduces the future income of the middle class. Theoretical motivations are rooted in the established correlation between inequality and reduced social mobility, as captured by the "Great Gatsby Curve," and we build on empirical findings that educational attainment is a critical factor in this dynamic. Literature highlights that countries with high-income inequality often have a higher private investment in education and lower public education expenditure, which is correlated with a larger proportion of students in private schools. This observation suggests that the dual education system plays a role in sustaining income inequality and reducing intergenerational mobility. In my model, we assume that private education is of higher quality due to productivity differences, leading parents with sufficient resources to invest in it over public education, thereby enhancing their offspring's human capital and economic prospects. However, we introduce a governmental role in funding compulsory public education, hypothesizing that increased public spending may disproportionately benefit lower-income groups while potentially disadvantaging the middle class.

Chapter 3, co-authored with Calin G. Arcalean and Ioana C. Schiopu, examines public education funding across U.S. school districts from 2005 to 2019, analyzing the relationship between income inequality and local funding patterns. Using a comprehensive dataset, we document that higher inequality correlates with reduced local public education funding per pupil, a trend primarily driven by poorer districts. In contrast, wealthier districts exhibit a neutral or slightly positive response, likely due to a more resilient tax base. To capture the causal effects of inequality on redistribution, we employ instrumental variables derived from synthetic, counterfactual income distributions, using exogenous shifts in national income as an instrument. This approach mitigates potential endogeneity due to demographic shifts and migration across districts.

Our results consistently show that higher income dispersion negatively impacts local funding, a robust pattern across various inequality measures (Gini coefficient, standard deviation of income, and coefficient of variation). These findings have significant policy implications. They suggest that public education funding mechanisms may inadvertently amplify existing inequalities, especially where local revenue sources dominate. Poorer districts facing higher inequality experience compounded disadvantages, as both local funding per pupil declines and state or federal contributions fail to fully offset these reductions. Our study highlights the importance of creating fair funding policies that consider how inequality affects districts with different income levels differently. This supports the idea that inequality has varied impacts on the provision of public goods.

In Chapter 4, I examine the evolving marriage trends in the United States from 2000 to 2021, focusing on two economic factors: changes in wage structures and working hours. Traditional analyses have often attributed declining marriage rates to shifts in gender roles and income disparities. However, recent data indicates that time availability, particularly working hours, may play a substantial role in marriage decisions. Using IPUMS data, I highlight a notable reduction in work hours for men across income groups, contrasted with a slight increase for women. These findings underscore that time, alongside financial considerations, is increasingly central to household formation. The study documents four key findings: (1) a narrowing gender wage gap alongside widening income disparities within gender groups; (2) a decrease in men's work hours, with a slight increase in women's; (3) a rise in positive assortative mating by income, seen in the increase of high-earning couples; and (4) contrasting marriage trends across gender and income groups, with high-income women showing increased marriage rates relative to their lower-income counterparts. To analyze these patterns, I develop a general equilibrium marriage model incorporating exogenous shifts in wages and working hours. By calibrating the model to 2000 data and validating it against 2021 trends, I conduct a counterfactual analysis of wage and time effects on marriage rates. The results reveal that while wage structure changes contribute to the general decline in marriage rates, shifts in work hours account for the increased marriage rates among high-income women. Specifically, reduced work hours among men appear to make marriage more appealing to high-income women by suggesting greater time availability for household activities.

2 Does public education hurt the middle class?

2.1 Introduction

Income inequality is correlated with inter-generational mobility. Countries with the highest income inequality are the least socially mobile. This negative relation between inequality and social mobility is known as the "Great Gatsby Curve" (GGC) (Ermisch et al.,2012). There is substantial research on the drivers that might explain the GGC.[1] In this regard, Jerrim and Macmillan (2015) finds that inter-generational mobility is driven in all countries by educational attainment. In addition, the negative link between income inequality and social mobility is stronger in more unequal countries. Jerrim and Macmillan (2015) provide empirical evidence suggesting that private investment in education, compared to public investment, is larger in more unequal countries. Moreover, these countries spend less on public education and have a higher proportion of children attending private schools or using private tutors. This might be due to the fact that public education has a poorer quality than private education, which is well documented in the literature.[2]

The main goal of this paper is to propose a theoretical framework to explain increased inequality and polarization due to segregation in education and analyze the short-run effect of public education on social classes, focusing on the middle class. This purpose is motivated by two reasons. The first is that a large and strong middle class spurs growth and reduces inequality.[3] The second is the existence of policies the government can take, which has been known to hurt the middle class.[4] Consequently,

[1]Check survey on the GGC by Durlauf et al. (2022).

[2]Coleman et al. (1982); Hanushek (1986); Psacharopoulos (1987); Chubb and Moe (1990); Jimenez et al. (1991); Neal (1997); Bedi and Garg (2000); Stevans and Sessions (2000) ; Mizzala et al. (2002); Bettinger (2005); Opdenakker and Van Damme (2006) ;Azimi et al. (2023); Crawfurd et al. (2023).

[3]Easterly (2001) provides empirical evidence that a middle-class consensus determines development outcomes and explains inequality.

[4]Simula and Trannoy (2010) finds that taxation represents a "curse" on the middle class when the government is Rawlsian. As per this paper, when taxes increase, the rich population migrates to other countries with lower tax rates. The middle class, on the other hand, which represents

the government should support and promote policies strengthening the middle class. Our paper provides new insights on how can government intervention in public education be a "curse" on the middle class under specific considerations.

The literature has paid tremendous attention to the role of education in the transmission of (dis)advantage across generations and inter-generational mobility. We contribute to two strands of the literature. The first one examines the role of wealth distribution in explaining inequality through investment in human capital. Galor and Zeira (1993) shows that when investment in education is indivisible, namely characterized by a technological non-convexity, the poorest individuals can acquire education only if they borrow. However, if capital market imperfections create borrowing constraints, low-income individuals are excluded from education, making upward social mobility unattainable. In contrast, rich individuals who inherit a large initial wealth have better access to investment in human capital without the need to borrow. As a result, initial wealth distribution persists and affects the rate of growth and inequality through its impact on the aggregate stock of human capital. Alonso-Carrera et al. (2012) contributes to this line of research by studying the impact of the joint distribution of bequest and human capital as well as fiscal policy on the persistence of inequality in the long run. I differ from this literature by introducing convex human capital technologies. Parents can either invest in public or private education. However, the choice between the type of education is discrete. This allows for the existence of social classes without the disruption of the non-convexity assumption.

The second studies the interaction between inequality and education choice. This literature is identified by the static analysis of education choice (Glomm and Ravikumar, 1998; Hoyt and Lee, 1998; de la Croix and Doepke, 2009a, Arcalean and Schiopu, 2015), and the dynamic inequality analysis in a given education regime, either public or private (e.g., Glomm and Ravikumar, 1992; Benabou, 2000; de la Croix and Doepke, 2004). I build on this literature to see particularly how government intervention in an economy with a dual education system has a heterogeneous impact on the education choice of different income groups.[5]

Our paper follows this line of research and contributes to it by introducing a dual education system consisting of public schools financed by the government and private schools, which are assumed to have better productivity than public education. We use an overlapping generations model representing a small open economy in which individuals live for three periods. In the first period, a young individual accumulates

the richer among those who are not rich enough to leave the country, incurs the larger part of the deadweight loss of taxation.

[5]To model a dual education we follow Brotherhood and Delalibera (2019). Other papers using similar education technology include Restuccia and Urrutia (2004).

human capital by acquiring an education financed by his parent. In the second period, he works, supplies labor, and chooses to educate his offspring in a private or public school. In the third period, he retires. Each parent has one child. As such, we assume no population growth.

In this overlapping generations model, we assume that parents are altruistic and care about the future income of their offspring. Individuals contribute to their children's future income by either giving a physical bequest, investing in education, or both. Consequently, agents derive utility from consumption and the transfers to children. We differ from the literature in three aspects. First, we consider a dual education system in which private schools have better quality in comparison to their public counterparts. Second, we assume a compulsory education system in which the government finances public schools. Finally, we assume that both private and public education technologies are convex.

The model generates four social classes that differ in the quantity in the quantity and types of transfers they provide to their children. Depending on the parametric conditions, the economy can exhibit different scenarios. We focus on an economy that is comparable to our empirical data and features three social classes: a poor class that invests in public education and does not provide bequests, a middle class that invests in private education and does not provide bequests, and a rich class that invests in private education and provides bequests. Our findings indicate that depending on specific parameter specifications, government expenditure on education affects the size of these social classes and their subsequent transfers. Specifically, we show that an increase in government spending reduces the middle class and increases the size of the poor class. Moreover, while increased public spending boosts the future income of children from lower-income families attending public schools, it can lower the income of some children from middle-class families by prompting their parents to switch from private to public schools. Improving the productivity of public education has a similar effect when the assumption of private education's superiority is maintained. However, if this assumption is relaxed, inequality decreases without hollowing out the middle class. Regarding taxation, we find that it reduces the optimal investment in education and decreases the future income of rich children, even though bequests increase to compensate for this decrease. It also improves the future income of some middle-class children whose parents choose to opt out of public schools.

The paper is presented as follows. Section 2 gives empirical motivation. Section 3 explains the proposed models and the assumptions upon which the analysis is built. Section 4 shows the equilibrium by solving the problem faced by individuals in this economy. In section 5, we study the effects of governmental intervention in public education on the middle class. Finally, in section 6, we conclude.

2.2 Empirical motivation

Schettino and Khan (2020) finds that the impoverishment of the middle-income class between the years 2000 and 2014, in reality, started in the 1980s and accelerated as time passed. The main premise of this paper is that government spending on education might be a good reason to explain this phenomenon through its impact on parental decisions over types of schools. We argue that family background and, specifically, educational expenditure is a good predictor of a child's future income. Rich families can send their kids to the most prestigious and the best schools in comparison to poor kids (Skiba et al., 2008). Affluent parents can also invest more in their children's preschool education and tap better early-age educational resources, as well as spending more on after-school training (Fan et al., 2020). Carnevale et al. (2019) assert that in the United States, a kid from a high-income family with low scores in kindergarten has a 70 percent chance of getting a college degree and an entry-level job. In contrast, a kid from a low-income family with high scores in kindergarten has only a 30 percent chance.

What is not clear is the extent to which government spending on education influences middle-class families to choose the best educational option for their kids. In this context, we use US data to evaluate the effect of government spending on the education decisions of middle-class families. We take the IPUMS American Community Surveys (ACS) from the year 2000 to 2019 and Public Elementary–Secondary Education Finance Data from the United states census bureau for the same period.[6] Using this data, we check if the parents' decision over public or private education is elastic to per capita public spending based on their income group. That is, we see the probability of parents opting out of private education when the government spending on education increases. It should be noted that in the theoretical model, we allow all individuals to "privately" invest in the education of their children based on income. This detail is mirrored in the data by the fact that private education is not exclusive to a specific group, although it is most common among the richer social classes. Our empirical model is as follows:

$$Private_i = \beta_0 Inc_i + \beta_1 Exp_i + \beta_2 \sum_{i=1}^{50} S_i + \beta_3 \sum_{j=20}^{65} A_{ji} + \beta_4 \; T_t + \epsilon$$

where $Private_i$ is a dummy variable set to take 1 if the household has at least one child in private school, Inc_i is the log of household income, Exp_i represents public expenditure per capita on primary and secondary education for household i, $State_i$ is a state dummy, and Age_{ji} is a dummy representing the age of the household head.

[6]We aggregate the variable to household level a la de la Croix and Doepke (2009a).

More controls are added to the main specification as a robustness check. These include household education and a vector of racial dummies.[7]

Estimating the regression using a Logit model, we test if the probability of choosing private education for individuals born in families whose income falls between the second and third quartiles of the income distribution reduces when per capita public expenditure in education increases. We control for state dummies to have the net impact of government spending. This is because controlling for the per capita spending alone could capture the effect of this variable on parents' choice through the channel of public education quality.

Results are summarized in Table (2.1). The estimation results indicate that an increase in per capita education spending is associated with a significant decrease in the likelihood of middle-class families choosing private schooling. In contrast, for poor and rich households, changes in per capita government spending do not show a statistically significant relationship with the probability of enrolling in private schools. We then account for household education and racial composition as robustness checks, and we find virtually the same results. Another robustness check is added in the Appendix. It runs the estimation with a metropolitan dummy as a control and uses the total public expenditure instead of primary and secondary public expenditure.

In the remainder of the paper, we build a theoretical model explaining the result we find and the possible consequences this might have on the composition of classes, their optimal decisions, and future inequality.

2.3 The model

We assume a small open economy populated by overlapping generations of individuals who live for three periods. In the first period, a young individual does not consume nor work, he accumulates human capital by attending school, which can be either public or private depending on his parent's decision. In the second period, he works, supplies labor, saves, and chooses between educating his offspring in a public or a private school. In the third period, he retires and allocates his savings between consumption and bequest. Each individual has one child at the beginning of the second period. Hence, there is no population growth. We assume a continuum of adult individuals of constant mass N.[8] Following the convention, we define generation t as the generation whose individuals are adult in period t

Agents derive utility from consumption in the second period t, consumption in the third period $t+1$, and their contribution to the lifetime income of their offspring.

[7] the racial dummies are for white, Hispanic, Pacific, Black, Asian, and other races.

[8] As there is no population growth, N is constant throughout

Table 2.1: Choice of Private Schooling on public spending per capita

	Middle class	Middle class	Rich	Rich	Poor	Poor
	(1)	(2)	(3)	(4)	(5)	(6)
Pub spending per capita	-0.398***	-0.375***	-0.108	-0.078	-0.208	-0.194
	(-4.780)	(-4.243)	(-0.912)	(-0.693)	(-1.019)	(-0.905)
Household Income	0.697***	0.418***	0.873***	0.796***	0.007	-0.011
	(15.237)	(14.667)	(30.085)	(28.389)	(0.594)	(-1.560)
Household educ	-	0.690***	-	0.638***	-	0.769***
		(24.193)		(15.687)		(20.575)
White	-	0.286***	-	0.237***	-	0.451***
		(6.571)		(6.112)		(8.177)
Hispanic	-	-0.290***	-	0.046	-	-0.449***
		(-3.864)		(0.516)		(-5.679)
Pacific	-	-0.095*	-	0.028	-	-0.486***
		(-1.700)		(0.485)		(-3.238)
Black	-	-0.246***	-	-0.041	-	-0.311***
		(-4.488)		(-0.696)		(-3.719)
Asian	-	0.051	-	0.061	-	0.052
		(1.343)		(1.569)		(0.504)
Other	-	0.220***	-	0.254**	-	0.242**
		(4.259)		(2.513)		(2.250)
Observations	2633037	2633037	1318289	1318289	1287456	1287456
Time fixed effects	✓	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is a dummy that takes 1 if the household has at least one child in private school. The primary/secondary public expenditure per capita and household income are expressed in logs of constant 2019 dollars. Covariates include household income and race dummies as well as time and age-fixed effects. The poor and rich groups represent the first and last quartiles, respectively. The middle class represents the second and third quartiles. For data sources and summary statistics see Appendix. Standard errors are clustered at the state level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Adult individuals contribute to their children's future income by either giving physical bequests, investing in their education, or both. Hence, the utility function of an individual from dynasty i and generation t is as follows:

$$U_t^i = \ln c_t^i + \rho \ln d_{t+1}^i + \beta \ln I_{t+1}^i \tag{2.1}$$

where ρ and β are strictly positive parameters capturing the temporal discount factor and the intensity of altruism respectively. c_t^i is consumption in the second period, whereas d_{t+1}^i is the consumption in the third period. I_{t+1}^i is the after-tax contribution to the children's lifetime income. A parent contribution to the income of an individual from dynasty i and generation $t+1$ is represented by the following:

$$I_{t+1}^i = (1-\tau)w_{t+1}h_{t+1}^i + b_{t+1}^i \tag{2.2}$$

where w_{t+1} is the wage per efficiency unit of labor at period $t+1$, $\tau \in [0,1)$ is the tax rate on labor income, h_{t+1}^i is the supplied labor efficiency units from dynasty i and generation $t+1$, and b_{t+1}^i is the bequest given to an individual belonging to dynasty i and generation $t+1$. Since we assume that a parent only has one descendant, the lifetime income of individuals from the same dynasty but different generation could differ if the contribution they receive from their respective parents is different. If the transfer received, on the other hand, is identical, then naturally, the lifetime income would be the same. For individuals belonging to the same generation but different dynasties, the transfers they receive would only be similar if the initial endowments of their respective dynasties were the same. Equation (2.2) has significant implications for intergenerational mobility, as variations in either the net labor income or the bequest can lead to divergent lifetime incomes among individuals affecting their economic prospects and social mobility.

In this model, we extend the framework established in Alonso-Carrera et al. (2012), introducing a dual education system as an additional variable to capture heterogeneity in educational choices. Specifically, we assume that a child's human capital is determined by the parental choice between public and private education, the subsequent investment in their education, and government spending. Let e_t^i denote the investment in children's education of an adult individual of dynasty i and generation t. We also assume that education is compulsory in this economy; that is, all young individuals receive an education. The level of human capital in period $t+1$ for an individual from dynasty i born in period t is determined by the following education technology, as introduced in Brotherhood and Delalibera (2019):

$$h_{t+1}^i = \begin{cases} \alpha(g + e_t^i)^\psi & \text{Public education} \\ (e_t^i)^\psi & \text{Private education} \end{cases} \tag{2.3}$$

where α is the parameter capturing the quality of public education. We assume that private education has a comparably better quality than public education. Hence, the parameter α is set such that $\alpha \in (0, 1)$. Parents who opt for public education might still have a positive investment in education e (e.g., purchasing books, hiring private tutors, etc.) together with g representing the per capita government's spending on public education. We assume that ψ is strictly less than one. This implies that human capital exhibits decreasing returns to education, which eliminates the possibility of sustained growth from our model.

Parents in this model face a trade-off when deciding between public and private education for their children. Opting for public education allows parents to utilize government funding (g). Still, parents can invest e_t^i to enhance their child's human capital. This investment may include supplementary educational resources, extracurricular activities, or private tutoring to compensate for the lower baseline quality. On the other hand, choosing private education typically requires a higher personal investment (e_t^i) but offers superior educational outcomes, thanks to the inherently higher quality of private institutions. This choice may lead to greater immediate costs but can potentially yield higher human capital (h_{t+1}^i) for the child. This decision is influenced by their current wealth, as wealthier parents may prefer the higher but costlier private education to maximize their child's human capital, while less affluent parents might rely more heavily on public education supplemented with limited or no personal investments.

In this economy, there is a single commodity that could either be consumed or invested, and investment made by adult individuals can either be in physical or in human capital. That is, the income of adult individuals, which is comprised of after-tax wage earnings and inheritance, is distributed between consumption, investment in the child's education, and saving. The budget constraint faced by an adult individual from dynasty i and generation t is then:

$$(1-\tau)w_t h_t^i + b_t^i = c_t^i + s_t^i + e_t^i \tag{2.4}$$

with s_t^i, c_t^i, and e_t^i representing the adult individual's savings consumption individual, and the amount he chooses to invest in his child's education. When an individual is in his third and last period, he receives a return on his saving, which is devoted proportionally to his consumption and bequest for his offspring. Therefore, the budget constraint for an individual in the third period is:

$$R_{t+1} s_t^i = d_{t+1}^i + b_{t+1}^i \tag{2.5}$$

where R_{t+1} is the gross rate of return on saving s_t^i, i.e., $R_{t+1} = 1 + r_{t+1}$. The return on savings is used by the old individual in $t+1$ to be consumed and given as a bequest to his offspring.

By combining (2.4) and (2.5), we have the following intertemporal budget constraint:

$$(1-\tau)w_t h_t^i + b_t^i = c_t^i + \frac{d_{t+1}^i + b_{t+1}^i}{R_{t+1}} + e_t^i \tag{2.6}$$

In this economy, we impose borrowing constraints on parents. By doing so, we avoid that parents borrow to educate their descendants. Becker and Tomes (1986) and Galor and Zeira (1993) have shown that if borrowing is constrained, education introduces intergenerational income persistence when parental investment in education depends only on parental income. When credit markets are perfect, the amount invested in the child's education is independent of parental income. Therefore, a borrowing constraint is necessary to explain intergenerational income persistence. This assumption is captured by the following condition:

$$b_{t+1}^i \geq 0 \tag{2.7}$$

(2.5) and (2.7) implies that $s_t^i \geq 0$; and, hence, borrowing is not possible.

Recall that the model is based on a small open economy. The interest rate, in this framework, is exogenously set in the international capital market such that $r = r^*$, where r^* is the world interest rate. We assume that the good of this economy is produced by means of a production function displaying constant returns to scale on physical and human capital. Moreover, let us assume that the stock of physical capital fully depreciates after one period. As such, the firm's technology can be written as follows:

$$Y_t = F(K_t, H_t) \tag{2.8}$$

where $H_t = \sum_{i=1}^{N} h_i^t$ is the total supply of efficiency units of labor in period t determined according to the education technology in (2.3). We can rewrite the production function as follows:

$$y_t = f(k_t) \tag{2.9}$$

where $y_t = \frac{Y_t}{L_t h_t}$ and $k_t = \frac{K_t}{L_t h_t}$. It should be noted that in this economy, the firms choose the ratio of physical to human capital in a manner consistent with their competitive behavior. That is, the firms' decisions are made such that the marginal productivity equals rental prices of physical and human capital. Based on the assumption of free mobility of physical capital, the ratio of physical to human capital ($\frac{K}{Lh}$) is constant (as $r^* = f'(\frac{K}{Lh})$). Therefore, the wage per efficiency unit of labor in equilibrium is set such that $w = f(\frac{K}{Lh}) - \frac{K}{Lh} f'(\frac{K}{Lh})$. Note that w is constant. Consequently, $w_t^i = w h_t^i$ for all t.

In this model, we assume that the government imposes solely a tax on labour income and spends the revenue to finance public education and on unproductive government spending, G_t^u. Denoting N the total population, the total government spending on public education can be written as the following:

$$G^E = gN\kappa \tag{2.10}$$

where $\kappa > 1$ is a parameter capturing the government inefficiency or bureaucratic cost. If one unit is devoted to public education by the government, only $1/\kappa$ would be effectively spent to achieve that purpose. g is, as defined in (3), the per capita government's allocation for public education.
The government faces a balanced budget constraint in each period. The total government spending G, is subject to the following condition at period t:

$$G_t = G_t^u + G^E = \int_0^N \tau w h_t^i di \tag{2.11}$$

In the government budget constraint, we are implicitly assuming that both the per capita spending on education $\bar{g}$, and tax τ, are constant and exogenous.

2.4 Individuals decisions

In this section, we address the optimization problem faced by an adult individual from dynasty i and generation t, who seeks to maximize their utility as defined in equation (2.1). The individual makes two distinct types of decisions: decisions on continuous variables and a discrete decision between two education systems. The continuous variables include consumption in the second period (c_t^i), consumption in the third period (d_{t+1}^i), the amount invested in their child's education (e_t^i), and the bequest to their child (b_{t+1}^i); and the discrete decision involves choosing between enrolling their child in public or private education. The optimization is subject to constraints (2.2), (2.3), (2.6), and (2.7) To effectively solve this problem, we adopt a two-step strategy.

First, for each possible choice of the education system (*public* or *private*) we solve the optimization problem with respect to the continuous variables $\{c_t^i, d_{t+1}^i, e_t^i, b_{t+1}^i\}$. This involves setting up the Lagrangian for each scenario and deriving the first-order conditions to determine the optimal levels of consumption, investment in education, savings, and bequest. During this step, the state variables, namely the inherited bequest b_t^i and the current human capital h_t^i, are treated as given.

Second, after obtaining the optimal solutions for the continuous variables under both public and private education scenarios, we compare the resulting utilities to determine which education system choice maximizes the individual's utility. This comparison allows us to identify whether the individual opts for public or private education based on his total income. Using these indirect utilities, we compute the threshold income at which the individual is indifferent between choosing public or private education. We further discuss the second step later in the paper.

Solving the first step of the problem, we obtain the following optimality conditions (the first-order conditions are computed in Appendix A):

$$c_t^i = \frac{(1-\tau)wh_t^i + b_t^i - \frac{b_{t+1}^i}{R} - e_t^i}{1+\rho} \tag{2.12}$$

$$d_{t+1}^i = c_t^i \rho R \tag{2.13}$$

and,

$$\frac{\beta}{I_{t+1}^i} \leq \frac{1+\rho}{R\left((1-\tau)wh_t^i + b_t^i - \frac{b_{t+1}^i}{R} - e_t^i\right)} \tag{2.14}$$

Equation (2.12) represents the optimal amount of consumption of an adult individual belonging to dynasty i and generation t, whereas equation (2.13) characterizes his optimal allocation of consumption along his lifetime. Equation (2.14) identifies the optimal amount of bequest this individual gives to his direct descendant. and it holds with equality when b_{t+1}^i is non binding, i.e., $b_{t+1}^i > 0$. For the sake of clarity, we can write (2.14) as the following:

$$\frac{\beta}{I_{t+1}^i} \leq \frac{1}{Rc_t^i}$$

Note that the left-hand side of (2.14) is the marginal utility gain received by an individual from increasing the amount of bequest b_{t+1}^i given to his child. The right-hand side, on the other hand, represents the marginal utility loss resulting from the decrease in his lifetime consumption because of an increase in the amount of bequest transferred to his offspring. Consequently, condition (2.14) ensures that when the non-negativity constraint on bequest is non-binding, there is no marginal variation in the utility of parents resulting from giving a larger amount of bequest to their children.

Substituting (2.12) and (2.13) in the intertemporal budget constraint in (2.6), we get the optimal amount of saving s_t^i as a function of h_t^i, b_t^i, e_t^i, and b_{t+1}^i. The latter variables represents the amount of intergenerational transfers. The optimal s_t^i obtained is as follows:

$$s_t^i = \frac{\rho\left((1-\tau)wh_t^i + b_t^i - e_t^i\right) + \frac{b_{t+1}^i}{R}}{1+\rho} \tag{2.15}$$

When the constraint (2.7) is non-binding, it is possible to compute the optimal amount of bequest given to the offspring from (2.14), which is characterized by the following equation:

$$\begin{aligned} b^i_{t+1} &\equiv B(h^i_t, b^i_t, e^i_t) \\ &= \frac{\beta R\left((1-\tau)wh^i_t + b^i_t - e^i_t\right)}{1+\rho+\beta} - \frac{\left((1-\tau)wh^i_{t+1}\right)(1+\rho)}{1+\rho+\beta} \end{aligned} \tag{2.16}$$

The optimal amount of bequest b^i_{t+1}, as specified in the equation, depends positively on the individual's endowments, h^i_t and b^i_t, and negatively on his investment in the education of his direct descendent, e^i_t.

Regarding the investment in education, parents are faced with two decisions. First, they choose between the two types of schooling systems. Contingent on this choice, children will acquire human capital as defined in (3). Second, adult individuals decide how much to invest in education, depending on the education system initially chosen. Note that the optimal levels of investment in education also vary depending on whether parents choose to make a bequest. All in all, a parent chooses public or private education, then decides how much to invest in his kid's education subject to his decision over bequest. Let us denote $\bar{e}_j$ when the constraint on bequest is non binding, i.e. $b^i_{t+1} > 0$, and $\hat{e}_j$ when the constraint is binding, i.e. $b^i_{t+1} = 0$. Using the first-order conditions in Appendix A, we obtain the following optimal levels of education investment when $b^i_{t+1} > 0$:

$$\tilde{e}_{pu} = \left(\frac{(1-\tau)w\alpha\psi}{R}\right)^{\frac{1}{1-\psi}} - g \tag{2.17}$$

and,

$$\tilde{e}_{pr} = \left(\frac{(1-\tau)w\psi}{R}\right)^{\frac{1}{1-\psi}} \tag{2.18}$$

with $\tilde{e}_{pr}$ and $\tilde{e}_{pu}$ representing the optimal level of investment in education in private and public schooling systems, respectively. The optimal investment in both public and private education does not depend on individual factors, including parental income. Moreover, investing in public education is less costly in comparison with private education ($\tilde{e}_{pu} \leq \tilde{e}_{pr}$). Hence, choosing private over public education when (2.7) is non-binding results in higher human capital for kids.[9] Substituting (2.17) and (2.18) in (2.16), we have two optimal amounts of bequest depending on the type of education parents choose for their direct decedents, i.e.:

[9]this result can be easily proven by substituting (2.17) and (2.18) in (2.3)

$$
\begin{aligned}
b^i_{pu,t+1} &\equiv B(I^i_t, \tilde{e}_{pu}) \\
&= \frac{\beta R\left(I^i_t - \tilde{e}_{pu}\right)}{1+\rho+\beta} - \frac{(1+\rho)(1-\tau)w\alpha(g+\tilde{e}_{pu})^{\psi}}{1+\rho+\beta}
\end{aligned}
\tag{2.19}
$$

$$
\begin{aligned}
b^i_{pr,t+1} &\equiv B(I^i_t, \tilde{e}_{pr}) \\
&= \frac{\beta R\left(I^i_t - \tilde{e}_{pr}\right)}{1+\rho+\beta} - \frac{(1+\rho)((1-\tau)w(\tilde{e}_{pr})^{\psi})}{1+\rho+\beta}
\end{aligned}
\tag{2.20}
$$

where $b^i_{pu,t+1}$ and $b^i_{pr,t+1}$ are the optimal amounts of bequest when parents invest in public and private education, respectively. Note that (2.17) and (2.18) represent the levels of education spending that maximizes the parent's utility the most, thus it is not profitable to educate(invest) more. What if parents can pay more than (2.17) and (2.18)? Any extra transfers will take the form of a bequest.
Conversely, when the non-negativity constraint on bequest is binding,i.e. $b^i_{t+1} = 0$, the amount of investment in public and private education is a function of the parent's endowments h^i_t and b^i_t. As shown in Appendix A, if an individual chooses not to leave a bequest and invest in private education, his educational investment would be:

$$
\hat{e}_{pr} = \frac{\beta\psi I^i_t}{1+\rho+\psi\beta} \tag{2.21}
$$

As for an individual choosing not to give a bequest and investing in public education, we find the following expression for the optimal amount of educational investment:

$$
\hat{e}_{pu} = \frac{\beta\psi I^i_t - (1+\rho)g}{1+\rho+\psi\beta} \tag{2.22}
$$

Similarly to the public and private investment levels, when the bequest is positive, we immediately obtain that $\hat{e}_{pu} < \hat{e}_{pr}$. The values of investment in education we have in (2.21) and (2.22) are the constraint non-optimal levels of education. As an individual's income increases, the constraint relaxes, and education increases. In this case, a parent can invest in education as long as the optimal levels of education $\tilde{e}_{pu}$ or $\tilde{e}_{pu}$ are not reached. Once $\hat{e}_{pu} = \tilde{e}_{pu}$ or $\hat{e}_{pr} = \tilde{e}_{pu}$, parents start giving bequest to their descendants.

Additionally, we impose a non-negativity constraint on both public education investment in (2.17) and (2.22) such that:

$$
\hat{e}_{pu} \geq 0 \quad \text{and} \quad \tilde{e}_{pu} \geq 0 \tag{2.23}
$$

This condition ensures that we do not have a negative investment in public education as government spending per capita increases. We discuss the implications of (2.23) later in the paper.

Based on the previous results, I summarize four distinct outcomes determined by the type of schooling chosen, the subsequent level of investment, and whether or not a bequest is provided to direct descendants. From here onward, we will categorize each choice as belonging to a specific social class. These social classes are distinguished as follows:

- *Social class 1:* does not give bequest and invests $\hat{e}_{pu}$ in public education.

- *Social class 2:* gives bequest $b^i_{pu,t+1}$ and invests $\tilde{e}_{pu}$ in public education.

- *Social class 3:* does not give bequest and invests $\hat{e}_{pr}$ in private education.

- *Social class 4:* gives bequest $b^i_{pr,t+1}$ and invests $\tilde{e}_{pr}$ in private education.

We move to the second step of the individual's optimization problem. In this step, the individual must make a discrete decision between two education systems: public and private, and decide whether to leave a bequest. Formally, the parent chooses one of the four social classes. For this, the decision hinges on which system yields the higher utility for the individual based on their total income and the utility outcomes from the previous step. Specifically, the individual compares the utility values obtained under each social class, using the derived optimal consumption and investment levels, to determine which class maximizes his overall lifetime utility. The decision rule is straightforward: the individual will choose a social class if the utility from this decision exceeds that of choosing the remaining classes.

To facilitate the comparison of utility outcomes across different social classes, we introduce the indirect utility functions for each social class. We denote these functions as V_{i} such that $i = \{1, 2, 3, 4\}$ representing social Classes 1 through 4, respectively. The indirect utilities represent the maximum attainable utility an individual can achieve under each specific set of choices regarding consumption, education investment, and bequests. By substituting the optimal values of the continuous variables derived in the first step into the individual's utility function (2.1), each indirect utility encapsulates the lifetime utility associated with a particular social class. This allows individuals to evaluate and compare the overall utility derived from selecting any of the four social classes given their economic endowments. The indirect utilities are:

$$V_1(I_t, \hat{e}_{pu}) \equiv (1+\rho+\psi\beta)\ln\left(\frac{I_t^i + g}{1+\rho+\psi\beta}\right) + \beta(1-\psi)\ln(\frac{\tilde{e}_{pu}+g}{\beta\psi}) + \rho\ln\rho R + \beta\ln\beta R \quad (2.24)$$

$$V_2(I_t, \tilde{e}_{pu}) \equiv (1+\rho+\beta)\ln\left(\frac{I_t^i - \tilde{e}_{pu} + \frac{g+\tilde{e}_{pu}}{\psi}}{1+\rho+\beta}\right) + \rho\ln\rho R + \beta\ln\beta R \quad (2.25)$$

$$V_3(I_t, \hat{e}_{pr}) \equiv (1+\rho+\psi\beta)\ln\left(\frac{I_t^i}{1+\rho+\psi\beta}\right) + \beta(1-\psi)\ln(\frac{\tilde{e}_{pr}}{\beta\psi}) + \rho\ln\rho R + \beta\ln\beta R \quad (2.26)$$

$$V_4(I_t, \tilde{e}_{pr}) \equiv (1+\rho+\beta)\ln\left(\frac{I_t^i + \frac{1-\psi}{\psi}\tilde{e}_{pr}}{1+\rho+\beta}\right) + \rho\ln\rho R + \beta\ln\beta R \quad (2.27)$$

Before comparing the indirect utilities across different social classes, we highlight two fundamental points that affect our analysis. First, in certain cases, choosing a particular social class might always be optimal regardless of the parent endowments. For example, if the indirect utility V_2 is strictly greater than V_1 for every I_t, then parents would always prefer to belong to social class 2. Second, some social classes are only feasible when specific conditions are met. This is essential to maintain the integrity of the model. Specifically, we consider four conditions that are a logical consequence of the non-negativity constraints on bequest (2.7) and public education investment (2.23). Two are an implication of the non-negativity constraint on bequests. and the remaining two are an implication of the public investment constraint.

Starting with the non-negativity constraint on bequests, recall that individuals in social classes 2 and 4 are characterized by making positive bequests to their descendants, satisfying the non-negativity constraint on bequests given by equation (2.7). That is, $b^i_{t+1,pu} > 0$ and $b^i_{t+1,pr} > 0$. This requirement implies the existence of income thresholds above which these positive bequests are possible, making social classes 2 and 4 feasible options. Using (2.19) and (2.20), we obtain I_2 which is the level of income at which $b^i_{t+1,pu} = 0$ and I_4 the level of income when $b^i_{t+1,pr} = 0$. These thresholds are calculated as follows:

$$I_2 = \frac{(1+\rho+\psi\beta)\tilde{e}_{pu} + (1+\rho)g}{\psi\beta} \quad (2.28)$$

and,

$$I_4 = \frac{(1+\rho+\psi\beta)\tilde{e}_{pr}}{\psi\beta}. \quad (2.29)$$

From (2.28) and (2.29), we can draw two key conclusions. First, I_4 is greater than I_2.[10]

Second, individuals with an income strictly higher than I_4 (I_2), can belong to social class 4 (2) because they can afford to make positive bequest while investing the

[10] we rearrange (2.28) as as $I_2 = \frac{(1+\rho+\psi\beta)(\tilde{e}_{pu}+g)}{\psi\beta} - g$. We know from (2.17) and (2.18) that $\tilde{e}_{pu} + g < \tilde{e}_{pu}$, therefore, $I_2 < I_4$

optimal amount in private (public) education. It is important to note that exceeding these income thresholds does not guarantee that parents will choose these social classes; rather, it makes these classes feasible options. In other words, the non-negativity constraint on bequest (2.7) limits the availability of certain social classes based on the individual's income. To illustrate, generic individuals with income $I_t^i < I_4$ can not belong to social class 2 or 4 because they are not willing to leave positive bequests while investing the optimal level in public education. Their feasible options are limited to social classes 1 and 3. Similarly, generic individuals with an income satisfying $I_2 \leq I_t^i < I_4$ can not belong to social class 4 for the same reasons. Their feasible options are social classes 1,2 and 3. As for individuals with income $I_t^i > I_4$, all four classes become feasible options. The intuition is as follows. Rich parents can choose any education system as long as it maximizes their respective utility, while poorer parents have limited choices.
Second, in addition to the borrowing constraint discussed earlier, there are two other conditions arising from the non-negativity constraint of public education investment. Specifically, Equations (2.17) and (2.22) must be positive (Condition 2.23). We analyze each implication of this constraint separately.
The inequality $\tilde{e}_{pu} \geq 0$ implies:

$$g \leq \bar{g} \quad \text{where} \quad \bar{g} = \left(\frac{(1-\tau)w\alpha\psi}{R} \right)^{\frac{1}{1-\psi}}. \tag{2.30}$$

This parametric condition sets a threshold for government spending that must not be exceeded for parents to invest a positive amount in public education. The intuition here is clear. If g is high enough, parents will not find it optimal to invest extra into the human capital of their direct descendants as a substitution effect exists between government spending and private investment in public education. This setup creates a conditional response based on the exogenously set value of g where the optimal investment amount is zero once government spending surpasses the threshold. We represent this outcome as follows:

$$\tilde{e}_{pu} = \begin{cases} \left(\frac{(1-\tau)w\alpha\psi}{R} \right)^{\frac{1}{1-\psi}} - g & \text{if} \quad g < \bar{g} \\ 0 & \text{if} \quad g \geq \bar{g} \end{cases} \tag{2.31}$$

The inequality $\hat{e}_{pu} \geq 0$ implies the existence of a threshold income level, denoted I_0, above which $\hat{e}_{pu}$ is guaranteed to be positive. By rearranging this inequality, we derive the threshold income I_0 as follows:

$$I_0 = \frac{(1+\rho)g}{\beta\psi}. \tag{2.32}$$

It is important to note that this threshold I_0 is lower than I_2. In fact, I_2 equals I_0 only when $\tilde{e}_{pu} = 0$. When these two thresholds are equal, we end up with an economy where parents choosing public education do not invest additional resources, as government spending alone satisfies the educational investment needs. This applies to both social classes 1 and 2.

To obtain the optimal class for each individual, we compute the thresholds at which individuals are indifferent between belonging to one social class or another. We get $V_4 > V_2$ if and only if the following condition is fulfilled:

$$(1-\tau)w\tilde{e}_{pr}^{\psi} - R\tilde{e}_{pr} > (1-\tau)w\alpha(\bar{g}+\tilde{e}_{pu})\psi - R\tilde{e}_{pu} \tag{2.33}$$

This equation represents the condition under which a parent will always choose to belong to social Class 4 (private education with bequests) over social Class 2 (public education with bequests). Specifically, the net return from investing an amount $\tilde{e}_{pr}$ in private education which is calculated as the offspring's future after-tax wage income from private education $(1-\tau)w\tilde{e}_{pr}^{\psi}$ minus the opportunity cost of this investment $R\tilde{e}_{pr}$, must exceed the net return from public education $((1-\tau)w\alpha(\bar{g}+\tilde{e}_{pu})^{\psi} - R\tilde{e}_{pu})$. By assuming that Condition (2.33) holds throughout the analysis, the model focuses on scenarios where higher-income individuals prioritize private education to maximize their child's future income, thereby reinforcing intergenerational income persistence.

we further have that $V_2 > V_1$ and $V_4 > V_3$ for specific income levels. These relationships are formally presented in the following proposition:

Proposition 1. *If $I_t > I_2$, then $V_2 > V_1$, and if $I_t > I_4$ then $V_4 > V_3$.*[11]

Summarizing, we establish that $V_4 > V_2 > V_1$ and $V_4 > V_3$, with I_2 and I_4 defining social classes 2 and 4, respectively. Additionally, the threshold I_0 identifies the subgroup of individuals who choose social class 1 but do not invest in public education. However, V_2 and V_1 are not comparable with V_3, necessitating the identification of additional threshold conditions to establish a ranking among them. This implies the existence of two additional thresholds. The first identifies the generic individual indifferent between belonging to social classes 1 and 3. The second identifies the generic individual indifferent between belonging to social classes 2 and 3. To find these thresholds, we solve the following equations:

$$V_1(I_t) = V_3(I_t) \tag{2.34}$$

$$V_2(I_t) = V_3(I_t) \tag{2.35}$$

Let's denote the solutions for (2.34) and (2.35) to be I_1 and I_3 respectively. Using the indirect utilities previously defined, we get that the solution for (2.34) is:

[11] Proof in Appendix

$$I_1 = \left(\frac{\alpha^{\frac{\beta}{1+\rho+\psi\beta}}}{1-\alpha^{\frac{\beta}{1+\rho+\psi\beta}}} \right) g \tag{2.36}$$

As for (2.35), there is no explicit solution. However, we can obtain insights by using an auxiliary function that we define as the difference between V_3 and V_2. Let us denote this function to be $\varphi(I_t)$, i.e.:

$$\begin{aligned} \varphi(I_t) = (1+\rho+\psi\beta)\ln\left(\frac{I_t^i}{1+\rho+\psi\beta}\right) + \beta(1-\psi)\ln\left(\frac{\tilde{e}_{pr}}{\beta\psi}\right) \\ -(1+\rho+\beta)\ln\left(\frac{I_t^i - \tilde{e}_{pu} + \frac{g+\tilde{e}_{pu}}{\psi}}{1+\rho+\beta}\right) \end{aligned} \tag{2.37}$$

Formally, there exists two solutions that solves $\varphi(I_t) = 0$ under certain condition as specified in the following proposition:

Proposition 2. *We obtain two solutions for the equation $\varphi(I_t) = 0$, denoted as I_3 and $\bar{I}_3$, when the following condition holds:*

$$(1-\psi)(\tilde{e}_{pr} - \tilde{e}_{pu}) > g \tag{2.38}$$

Assuming that (2.38) holds, implies the existence of the threshold I_3 that identifies social classes 3 and 2. The ranking of this threshold with regard to I_4 is:

Proposition 3. *If* (2.38) *holds, we get $I_3 < I_4 < \bar{I}_3$.*

Proposition (3) entails a very important corollary: $\bar{I}_3$ is irrelevant as long as it is greater than I_4. This is due to the fact that individuals with income greater than I_4, always choose to belong to social class 4 rather than social class 2 ($V_4 > V_2$).

To determine the optimal social class for each individual, we identify four income thresholds: I_1, I_2, I_3, and I_4. The other income threshold I_0 identifies the subgroup of individuals who belong to social class 1 and do not invest in public education. These thresholds represent the income levels at which individuals are indifferent between belonging to different social classes. Among these, certain thresholds can be systematically ranked. Specifically, we have $I_4 > I_2 > I_0$ and $I_4 > I_3$. However, not all thresholds are directly comparable due to the different constraints inherent in the model (i.e. I_1 and I_3 are not comparable to I_0 and I_2). The thresholds that cannot be sequentially ranked relative to one another give rise to distinct "scenario economies." Each scenario economy represents a unique set of parametric conditions and heterogeneous behaviors, reflecting different configurations of income thresholds. In the following section, I discuss the possible rankings of I_1 and I_3 with respect to I_0 and I_2, and I analyze the consequent implications on the existence and diversity of the aforementioned scenario economies.

It is easy to observe that when I_1 is feasible, I_3 is not, and vice-versa. If $I_1 > I_3$, all individuals with an income falling between these thresholds $I_1 \geq I_t \geq I_3$ would be indifferent between belonging to social classes 2 and 3 which is not possible, as V_2 is equal to V_3 only at I_3. If $I_3 > I_1$, all individuals with an income between I_3 and I_1 would be indifferent between belonging to social class 3 and 1 which is not possible. This implies that I_1 and I_3 do not exist simultaneously in the economy. That is, if we consider one of these two thresholds, we automatically disregard the other. This is also implied for different rankings of I_1 and I_3 with I_0 and I_2.[12]

Consequently, the model has three possible economies. We summarize the possible social group compositions in five economies:

- **Economy 1**: If $I_4 > I_2 > I_0 > I_1$, we get a three-class economy with social classes 1, 3, and 4. The thresholds separating the three groups are I_1 and I_4, respectively. In this economy, all parents in social class 1 invest in public education.

- **Economy 2**: If $I_4 > I_3 > I_2$, we get a four-class economy with social classes 1, 2, 3, and 4. The thresholds separating these groups are I_2, I_3, and I_4 respectively. In this economy, the threshold I_0 identifying social class 1 parents who do not invest in public education is feasible.

- **Economy 3**: If $I_4 > I_1 > I_2$, we get a three-class economy with social classes 1, 2, and 4. The thresholds separating these groups are I_2 and I_4. In this economy, the threshold I_0 is feasible.

- **Economy 4**: If $I_4 > I_2 > I_1 > I_0$, we get a three-class economy with social classes 1, 3, and 4. The thresholds separating these groups are I_1 and I_4. In this economy, the threshold I_0 is feasible.

- **Economy 5**: If $I_4 > I_2 > I_3 > I_0$, we get a three-class economy with social classes 2, 3, and 4. The thresholds separating these groups are I_3 and I_4. In this economy, the threshold I_0 is feasible.

[12] The ranking of I_1 with respect to I_0 and I_2, depends on α thresholds. We define α_1 and α_2 as the solutions for $I_0 = I_1$ and $I_1 = I_2$, respectively. We get:

$$\alpha_1 = \left(\frac{1+\rho}{1+\rho+\beta\psi}\right)^{\frac{1+\rho+\beta\psi}{\beta}} \quad \text{and} \quad \alpha_2 = \left(1 - \frac{g}{\bar{g}} \cdot \frac{\psi\beta}{1+\rho+\psi\beta}\right)^{\frac{1+\rho+\psi\beta}{\beta}}$$

We can rewrite $\alpha_2 = 1 - \frac{\psi\beta}{(1+\rho+\psi\beta)}$. Since $\frac{g}{\bar{g}} < 1$, we directly get that $\alpha_2 > \alpha_1$. At $g = \bar{g}$ (the case where public education investment in social class 1 is zero) $\alpha_2 = \alpha_1$. When $\alpha < \alpha_1$, we have $I_1 < I_0$. When $\alpha_1 < \alpha < \alpha_2$, we have $I_0 < I_1 < I_2$. When $\alpha > \alpha_2$, we have $I_2 < I_1$. One interesting conclusion is that when public school productivity is low more parents belong to social class 3 (private education), whereas the opposite happens when productivity is high and closer to 1.

For the remainder of this analysis, we will focus on Economy 1 as our primary scenario because it consists of three groups, making it more representative of the data. Additionally, the dynamics in Economy 1 are applicable to Economy 4. Economy 3 is excluded from the analysis as it lacks social class 3. An analysis of Economy 2, which is applicable to Economy 5, is included in the appendix for reference.

2.5 Government intervention in education and the middle class

In this section, we focus on the effect of government policies in public education on the middle class for both economies in the short run. Particularly, we center our analysis around the effect of the marginal change of three parameters on the thresholds I_1, I_3, I_2, and I_4. These parameters are the income tax τ, the per capita government spending on education g, and the quality of public education α. Note that changes in these parameters affect the short-term size of social groups and their respective income. We proceed by analyzing the impact of the parameter changes for each economy separately.

It is important to note that our analysis in this section is concentrated on the short-term effects of these policy changes. We examine how immediate adjustments in policy parameters influence the size and income of children born in the middle class and other social classes without delving into the long-term dynamics. A thorough investigation of the long-term effects would require analyzing the steady-state equilibrium of the economy, which is beyond the scope of this paper.

2.5.1 Economy 1

In this economy, we have three different groups identified by the threshold I_1 and I_4. Parents' decisions are made based on the level of their income and what social class they belong to. We summarize the three social classes in this economy as follows:

- *Poor:* Has an income below I_1, does not give bequest and invests $e^i_{pu,t}$ in public education (social class 1).

- *Middle class:* Has an income above I_1 but below I_4, does not give bequest and invests $e^i_{pr,t}$ in private education (social class 3).

- *Rich:* Has an income above I_4, gives bequest $b^i_{pr,t+1}$ and invests e^i_{pr} in private education (social class 4).

The marginal effect of government spending in public education g

In this subsection, we analyze how an increase in government spending on public education (g) affects the social class composition in Economy 1, with a particular focus on the middle class. Our objective is to understand the implications of higher public education funding on parents' educational choices and the subsequent impact on their children's human capital formation.

To begin, we examine the effect of an increase in g on the income threshold I_1, which separates the poor class from the middle class. By differentiating I_1 with respect to g, we obtain:

$$\frac{\partial I_1}{\partial g} = \frac{\alpha^{\frac{\beta}{1+\rho+\psi\beta}}}{1-\alpha^{\frac{\beta}{1+\rho+\psi\beta}}} > 0. \tag{2.39}$$

Since $\alpha \in (0,1)$, both the numerator and the denominator are positive, ensuring that the derivative is positive. This positive relationship indicates that as government spending on public education increases, the threshold I_1 rises. Individuals whose incomes were just above the previous I_1 now find themselves below the new, higher threshold, effectively expanding the size of the poor class. Notably, the threshold I_4, which separates the middle class from the rich class, remains unchanged because it does not depend on g. Therefore, the size of the rich class remains unaffected by changes in government spending on education and the income range defining the middle class narrows, leading to a shrinkage of the middle class.

The magnitude of the threshold shift is influenced by the quality of public education, represented by α. A higher α (closer to 1) amplifies the shift in I_1, resulting in a more significant contraction of the middle class. This outcome is intuitive: as public education becomes more productive, parents perceive greater value in public schooling, leading some to opt out of private education in favor of the improved public option.

Next, we explore the impact of increased government spending on educational investment and human capital levels for the different social classes. For the rich and middle classes, the levels of educational investment remain unchanged. This is evident from equations (2.18) and (2.21), which show that the optimal investments $\tilde{e}_{pr}$ and $\hat{e}_{pr}$ are independent of g. As a result, the human capital of the next generation within these classes is unaffected by changes in government spending.

In contrast, the poor class exhibits a different response. Parents in this class adjust their educational investment in reaction to changes in g. Differentiating $\hat{e}_{pu}$ with respect to g, we find:

$$\frac{\partial \hat{e}_{pu}}{\partial g} = -\frac{1+\rho}{1+\rho+\psi\beta} \in (-1,0). \tag{2.40}$$

This negative relationship implies that as government spending on education increases, parents in the poor class reduce their own investment $\hat{e}_{pu}$, but not by as much as the increase in g. This partial offset occurs because government spending and parental investment are imperfect substitutes in the human capital production function, as shown in equation (2.3). Despite reducing their own spending, the total educational resources available to their children ($g + \hat{e}_{pu}$) increase, leading to an improvement in the children's human capital and future income.

The reduction in parental investment allows parents in the poor class to reallocate resources toward their own lifetime consumption (c_t and d_{t+1}). The extent of this reallocation and the trade-off between investing in children's education versus personal consumption depend on two key parameters: the altruism parameter (β) and the concavity parameter of the human capital production function (ψ).

Specifically, when parents are less altruistic (lower β), they place greater emphasis on current consumption over investing in their children's future income. In this case, the reduction in $\hat{e}_{pu}$ approaches the full amount of the increase in g ($\left|\frac{\partial \hat{e}_{pu}}{\partial g}\right| \approx 1$). Conversely, if parents are more altruistic (higher β), they are less inclined to reduce their investment in education, resulting in a smaller decrease in $\hat{e}_{pu}$ ($\left|\frac{\partial \hat{e}_{pu}}{\partial g}\right| \approx 0$).

The concavity parameter ψ also plays a crucial role. A higher ψ indicates that the human capital production function is less concave, meaning that additional spending on education yields higher marginal returns. In such cases, parents are less likely to reduce their educational investment when g increases, as the benefits of additional investment are more pronounced.

An important result is the impact on middle-class individuals whose incomes are just above I_1 before the increase in g. As I_1 rises, some of these individuals find it optimal to switch from private to public education, effectively moving from the middle class to the poor class. This shift results in a decrease in their children's human capital, as they now receive lower-quality education (by the assumption that private education is superior) and benefit less from parental investment. Using (2.2), (2.3), (2.21) and (2.22), we can easily show that the income of children whose parents belong to social class 3 is higher than the income of children whose parents belong to social class 1.

From the parents' perspective, this change represents a net gain in utility. By opting for public education, they can reduce educational expenses and increase their own consumption.

These results are consistent with the evidence presented in Section 2, where it is shown that middle-class families tend to opt out of private education as government spending on public education increases. The theoretical framework developed here provides a rationalization for this behavior.

The Marginal Effect of Taxation (τ)

The impact of changes in taxation (τ) on different social classes is analyzed through its effect on income thresholds and investment decisions. Notably, the threshold I_1, which separates the poor from the middle class, remains unaffected by changes in taxation since it is independent of τ. Consequently, the size of the poor class remains constant, as does their level of investment in education.

However, the threshold I_4, which separates the middle class from the rich class, is influenced by changes in taxation. Differentiating I_4 with respect to τ yields:

$$\frac{\partial I_4}{\partial \tau} = \frac{-(1+\rho+\psi\beta)}{\psi\beta(1-\psi)(1-\tau)} \left(\frac{(1-\tau)w\psi}{R}\right)^{\frac{1}{1-\psi}} < 0 \tag{2.41}$$

This result indicates that, as taxes increase, I_4 decreases. The reduction in I_4 decreases the size of the middle class, as some individuals near the upper end of this class transition into the rich class. This shift is primarily driven by changes in the optimal level of private education investment for the rich, $\tilde{e}_{pr}$, which is negatively correlated with taxation:

$$\frac{\partial \tilde{e}_{pr}}{\partial \tau} = \frac{-1}{(1-\psi)(1-\tau)} \left(\frac{(1-\tau)w\psi}{R}\right)^{\frac{1}{1-\psi}} < 0 \tag{2.42}$$

An increase in taxes lowers the optimal private education investment $\tilde{e}_{pr}$, making it more affordable for a subset of middle-class individuals. These individuals, whose incomes were previously marginally below I_4, now find it optimal to invest at the rich class level. Consequently, they start exhibiting behavior typical of the rich class, such as making the optimal private education investment and providing bequests.

For individuals already classified as rich, an increase in taxation reduces their optimal level of private education investment, leading to a decrease in their children's human capital. To mitigate this effect, rich parents increase the bequests they leave to their children. This adjustment partially offsets the decline in the children's future income caused by the reduced investment in education. However, the compensatory bequest does not fully make up for the loss in human capital, eventually leading to a reduced income for the children. We can illustrate this effect by differentiating the child's income with respect to the tax rate τ:

$$\begin{aligned} \frac{\partial I_{t+1}}{\partial \tau} &= \frac{\partial h_{t+1}}{\partial \tau} + \frac{\partial b_{pr,t+1}}{\partial \tau}, \\ \text{where} \quad \frac{\partial h_{t+1}}{\partial \tau} &= \frac{-1}{(1-\psi)(1-\tau)} \tilde{e}_{pr}, \\ \text{and} \quad \frac{\partial b_{pr,t+1}}{\partial \tau} &= \left|\frac{\partial h_{t+1}}{\partial \tau}\right| \cdot \frac{1+\rho+\psi\beta}{1+\rho+\beta}. \end{aligned} \tag{2.43}$$

This demonstrates that while higher taxes decrease the children's human capital (since $\frac{\partial h_{t+1}}{\partial \tau} < 0$), parents increase their bequests (as $\frac{\partial b_{pr,t+1}}{\partial \tau} > 0$) to offset this effect. However, because the bequest does not fully compensate for the reduction in human capital, the children's total income still decreases.

The marginal effect of a change in quality α

An increase in the quality of public education (α) has a significant impact on the income threshold I_1 that separates the poor from the middle class. By differentiating I_1 with respect to α, we find:

$$\frac{\partial I_1}{\partial \alpha} = g \frac{\beta}{1+\rho+\psi\beta} \left(\frac{\alpha^{\frac{\beta}{1+\rho+\psi\beta}-1}}{(1-\alpha^{\frac{\beta}{1+\rho+\psi\beta}})^2} \right) \tag{2.44}$$

This result implies that an increase in the quality of public education leads to a reduction in the size of the middle class. Similar to the effect of increased government spending g, a marginal group within the middle class opts out of private schooling, preferring higher indirect utility over a better lifetime income for their children. From equation (2.44), we observe that the impact of the change in quality is more pronounced when parents place less importance on the future income of their descendants. That is, when β is smaller.

However, unlike the effect of increased government spending, an improvement in α enhances the productivity of both government and parental investment in education. Consequently, the optimal amount of investment in public education for adults belonging to the poor class, denoted as $\hat{e}_{pu}$, remains unchanged, and the lifetime income of their children increases. We show this by taking the derivative of the future income with respect to τ for social class 2:

$$\frac{\partial I_{t+1}}{\partial \tau} = (1-\tau)w(g+\hat{e}_{pu}) > 0 \tag{2.45}$$

2.6 Conclusion

In this paper, we empirically demonstrate that, in the U.S., the middle class's decision regarding the type of education for their children is more sensitive to changes in government spending than that of the poor and rich classes. Specifically, our findings show that increases in public education spending lead middle-class families to opt out of private schools more than other social classes. Motivated by this empirical evidence, we present a model of overlapping generations where parents care about the welfare of their direct descendants and contribute to their income by either investing in their education or giving them a direct transfer in the form bequest. We assume that parents can invest in public or private education depending on their stock of wealth. Moreover, we introduce a parameter α capturing the quality of public education and impose that it is strictly lower than one to establish the superiority of private education. In addition, we assume that individuals in this economy are credit-constrained, so they cannot borrow to invest in education. This implies that if a poor individual cannot afford private education, they would invest in public education or not invest at all.

We deduce that different social stratifications are possible and we focus on one case that we consider is particularly relevant. One with four classes and one with three classes. These classes are differentiated by their levels of human capital and bequest as well as subsequent transfers to their offsprings. We compute the income thresholds defining the different social classes. Parents choose an education system, invest in education and decide to leave bequest or not, depending on whether their endowed income is below or above the threshold computed. We use these thresholds to analyze the effect of three main policies on the distribution of individuals across the different classes. Particualrly, we analyse the impact of government spending on education, taxes, and productivity of the public education.

The model shows that higher public spending can unintentionally lower the future income of middle-class children by encouraging a shift from private to less productive public education. This outcome provides a rationale for the behaviors documented in our data, where such a transition reduces the size of the middle class by limiting opportunities for upward mobility. We also examine how policies like changes in income tax rates and improvements in public education quality affect the different social classes.

These findings highlight the complexity of education policy and its varied impacts on different social groups. Policymakers need to consider these differences to avoid unintended negative effects on the middle class, which is crucial for reducing inequality. Future research should explore ways to improve public education quality without disadvantaging the middle class and investigate the long-term effects of these educational choices on economic inequality and social mobility. This policy analysis could be done in a quantitative model that incorporates public and private education decisions and generates social classes.

Appendix A

A.1 First order conditions

We derive the optimal conditions on c_t^i, d_{t+1}^i, e_t^i, s_t^i and b_{t+1}^i. We maximize ((2.1)) with respect to $\{c_t^i, d_{t+1}^i, e_t^i, b_{t+1}^i\}$ subject to (2.4), (2.5), (2.6) and (2.7). As result we get the following Lagrangian with λ_t as the lagrangian multiplier:

$$\mathcal{L}_t = \ln c_t^i + \rho \ln d_{t+1}^i + \beta \ln I_{t+1}^i + \lambda_t \left((1-\tau) w h_t^i + b_t^i - c_t^i - \frac{d_{t+1}^i + b_{t+1}^i}{R_{t+1}} - e_t^i \right) \tag{2.46}$$

The first order conditions (FOC) of this problem are given by:

$$c_t^i = \frac{1}{\lambda_t} \tag{2.47}$$

$$d_{t+1}^i = \frac{\rho R}{\lambda_t} \tag{2.48}$$

$$\frac{\beta(1-\tau) w \alpha \psi (\bar{g} + e)^{\psi - 1}}{I_{t+1}^i} = \lambda_t \quad \text{For public education} \tag{2.49}$$

$$\frac{\beta(1-\tau) w \psi e^{\psi - 1}}{I_{t+1}^i} = \lambda_t \quad \text{For private education} \tag{2.50}$$

and

$$I_{t+1}^i \geq \frac{\beta R}{\lambda_t} \tag{2.51}$$

By substituting Equation (2.47) into Equation (2.6), we obtain the optimal level of c_t as follows:

$$I_t = c_t + \frac{\rho R c_t + b_{t+1}}{R} + e_t \tag{2.52a}$$

$$\Rightarrow c_t = \frac{I_t - \frac{b_{t+1}}{R} - e_t}{1+\rho} \tag{2.52b}$$

Optimal level of d_{t+1}:

$$d_{t+1} = \rho R c_t. \tag{2.53}$$

We know that when $b_{t+1} > 0$, (2.51) holds at equality. Thus,

$$I_t = \beta R c_t \tag{2.54a}$$

$$\Rightarrow (1-\tau)wh_{t+1} + b_{t+1} = \beta R \left(\frac{I_t - \frac{b_{t+1}}{R} - e_t}{1+\rho} \right) \tag{2.54b}$$

$$\Rightarrow b_{t+1} = \frac{\beta R(I_t - e_t) - (1+rho)(1-\tau)wh_{t+1}}{1+\rho+\beta}. \tag{2.54c}$$

Note that b_{t+1} has to be positive. This implies $\beta R(I_t - e_t) > (1+rho)(1-\tau)wh_{t+1}$.

Now we compute the optimal levels of education for public and private when $b_{t+1} = 0$ and $b_{t+1} > 0$. First, we have:

$$\frac{\partial \mathcal{L}_t}{\partial e_{pu}} = 0 \quad \Rightarrow \quad \frac{\beta(1-\tau)w\alpha\psi(g+e_{pu})^{\psi-1}}{I_{t+1}} = \lambda_t, \tag{2.55}$$

and

$$\frac{\partial \mathcal{L}_t}{\partial e_{pr}} = 0 \quad \Rightarrow \quad \frac{\beta(1-\tau)w\psi(e_{pr})^{\psi-1}}{I_{t+1}} = \lambda_t. \tag{2.56}$$

When $b_{t+1} > 0$ we use (2.51) with equality and substitute in the (2.55) and (2.56). We get:

$$\beta(1-\tau)w\alpha\psi(g+\tilde{e}_{pu})^{\psi-1} = \lambda_t \frac{\beta R}{\lambda_t} \tag{2.57a}$$

$$\beta(1-\tau)w\alpha\psi(g+\tilde{e}_{pu})^{\psi-1} = \beta R \tag{2.57b}$$

$$(g+\tilde{e}_{pu})^{\psi-1} = \frac{(1-\tau)w\alpha\psi}{R} \tag{2.57c}$$

$$g+\tilde{e}_{pu} = \left(\frac{(1-\tau)w\alpha\psi}{R} \right)^{\frac{1}{1-\psi}} \tag{2.57d}$$

$$\tilde{e}_{pu} = \left(\frac{(1-\tau)w\alpha\psi}{R} \right)^{\frac{1}{1-\psi}} - g \tag{2.57e}$$

$$\beta(1-\tau)w\psi(\tilde{e}_{pr})^{\psi-1} = \lambda_t \frac{\beta R}{\lambda_t} \tag{2.58a}$$

$$\beta(1-\tau)w\psi(\tilde{e}_{pr})^{\psi-1} = \beta R \tag{2.58b}$$

$$(\tilde{e}_{pr})^{\psi-1} = \frac{(1-\tau)w\psi}{R} \tag{2.58c}$$

$$\tilde{e}_{pr} = \left(\frac{(1-\tau)w\psi}{R} \right)^{\frac{1}{1-\psi}}. \tag{2.58d}$$

When $b_{t+1} = 0$, we use (2.54b) and substitute in (2.55) and (2.56), we get:

$$\frac{\beta(1-\tau)w\alpha\psi(g+\hat{e}_{pu})^{\psi-1}}{I_{t+1}} = \frac{1}{c_t} \tag{2.59a}$$

$$\frac{\beta(1-\tau)w\alpha\psi(g+\hat{e}_{pu})^{\psi-1}}{(1-\tau)w\alpha(g+\hat{e}_{pu})^{\psi}} = \frac{1+\rho}{I_t - \frac{b_{t+1}}{R} - \hat{e}_{pu}} \tag{2.59b}$$

$$\frac{\beta\psi}{g+\hat{e}_{pu}} = \frac{1+\rho}{I_t - \hat{e}_{pu}} \tag{2.59c}$$

$$\hat{e}_{pu} = \frac{\beta\psi I_t - (1+\rho)g}{1+\rho+\beta} \tag{2.59d}$$

$$\tag{2.59e}$$

$$\frac{\beta(1-\tau)w\psi(\hat{e}_{pr})^{\psi-1}}{I_{t+1}} = \frac{1}{c_t} \tag{2.60a}$$

$$\frac{\beta(1-\tau)w\psi(\hat{e}_{pr})^{\psi-1}}{(1-\tau)w(\hat{e}_{pr})^{\psi}} = \frac{1+\rho}{I_t - \frac{b_{t+1}}{R} - \hat{e}_{pr}} \tag{2.60b}$$

$$\frac{\beta\psi}{\hat{e}_{pr}} = \frac{1+\rho}{I_t - \hat{e}_{pr}} \tag{2.60c}$$

$$\hat{e}_{pr} = \frac{\beta\psi I_t}{1+\rho+\beta} \tag{2.60d}$$

A.2 Proof for proposition 1

Proof. We know that $V_4(I_t)$ is defined such that $I_t \in]I_4, +\infty[$. Computing the limit when I_t goes to I_4, we get :

$$\lim_{I_t \to I_4} V_4 = (1+\rho+\beta)\ln(\frac{\tilde{e}_{pr}}{\psi\beta}) + \rho\ln\rho R + \beta\ln\beta R$$

we know that $V_3(I_t)$ is defined over $\mathbb{R}^+$. Computing $V_3(I_4)$ we get:

$$V_3(I_4) = (1+\rho+\beta)\ln(\frac{\tilde{e}_{pr}}{\psi\beta}) + \rho\ln\rho R + \beta\ln\beta R$$

Notice that

$$\lim_{I_t \to I_4} V_4 = V_3(I_4)$$

We then compute the slope of both functions with respect of I_t, we get:

$$\frac{\partial V_4}{\partial I_t} = \frac{1+\rho+\beta}{I_t + (\frac{1-\psi}{\psi})\tilde{e}_{pr}} \quad \text{and,} \quad \frac{\partial V_3}{\partial I_t} = \frac{1+\rho+\psi\beta}{I_t}$$

It is easy to show that $\frac{\partial V_4}{\partial I_t} > \frac{\partial V_3}{\partial I_t}$ if and only if $I_t > I_4$. Consequently, $V_4 > V_3$ if $I_t > I_4$

Similarly,We know that $V_2(I_t)$ is defined such that $I_t \in]I_2, +\infty[$. Computing the limit when I_t goes to I_2, we get :

$$\lim_{I_t \to I_2} V_2 = (1+\rho+\psi\beta)\ln(\frac{\tilde{e}_{pu}+g}{\psi\beta}) + \rho\ln\rho R + \beta\ln\beta R$$

we know that $V_1(I_t)$ is defined over $\mathbb{R}^+$. Computing $V_1(I_2)$ we get:

$$V_1(I_2) = (1+\rho+\psi\beta)\ln(\frac{\tilde{e}_{pu}+g}{\psi\beta}) + \rho\ln\rho R + \beta\ln\beta R$$

Notice that

$$\lim_{I_t \to I_2} V_2 = V_1(I_2)$$

We then compute the slope of both functions with respect of I_t, we get:

$$\frac{\partial V_2}{\partial I_t} = \frac{1+\rho+\beta}{I_t - \tilde{e}_{pu} + \frac{g+\tilde{e}_{pu}}{\psi}} \quad \text{and,} \quad \frac{\partial V_1}{\partial I_t} = \frac{1+\rho+\psi\beta}{I_t+g}$$

It is easy to show that $\frac{\partial V_2}{\partial I_t} > \frac{\partial V_1}{\partial I_t}$ if and only if $I_t > I_2$. Consequently, $V_2 > V_1$ if $I_t > I_2$

All in all,

$$\begin{cases} I_t > I_4 \Leftrightarrow V_4 > V_3 \\ I_t > I_2 \Leftrightarrow V_2 > V_1 \end{cases} \tag{2.61}$$

□

A.3 Proof for proposition 2

Proof. Using (2.37) we compute the derivative with respect of I_t. We get :

$$\frac{\partial\varphi}{\partial I_t} = \frac{(1+\rho+\psi\beta)(I_t - \tilde{e}_{pu} + \frac{g+\tilde{e}_{pu}}{\psi}) - (1+\rho+\beta)I_t}{I_t(I_t - \tilde{e}_{pu} + \frac{g+\tilde{e}_{pu}}{\psi})}$$

We then solve $\frac{\partial\varphi}{\partial I_t} = 0$ to obtain the income at which the function φ reaches its maximum level. We call this income level I_m:

$$I_m = \frac{(1+\rho+\psi\beta)((1-\psi)\tilde{e}_{pu}+g)}{(1-\psi)\psi\beta}$$

As such, when $I_t < I_m$, $\varphi(I_t)$ is increasing, and when $I_t > I_m$, $\varphi(I_t)$ is decreasing. We can conclude with ease that the function $\varphi(I_t)$ will intersect twice with the horizontal axis if and only if there exists a range of I_t for which $\varphi(I_t)$ is strictly positive.

We compute $\varphi(I_m)$, and we get:

$$\varphi(I_m) = (1-\psi)\beta\ln\left(\frac{(1-\psi)\tilde{e}_{pr}}{(1-\psi)\tilde{e}_{pu}+g}\right)$$

If $\varphi(I_m) > 0$ there exists two points at which $\varphi(I_t)$ intersect with the horizontal axis, let us denote them I_3 and $\bar{I}_3$ such that $I_3 < \bar{I}_3$. This implies the following condition:

$$(1-\psi)(\tilde{e}_{pr} - \tilde{e}_{pu}) > g$$

If this condition is held with equality $\varphi(I_t)$ intersect with the horizontal axis in one point which is I_m. and if this condition is reversed $\varphi(I_t)$ does not intersect at all with with the horizontal axis. □

A.4 Proof for proposition 3

Proof. We know that the maximum of $\varphi(I_t)$ is:

$$I_m = \frac{(1+\rho+\psi\beta)((1-\psi)\tilde{e}_{pu}+g)}{(1-\psi)\psi\beta}$$

such that $I_3 < I_m < \bar{I}_3$ when (2.38) holds. Comparing I_4 from (2.29) with I_m, we easily get that:

$$I_4 > I_m \Leftrightarrow (1-\psi)(\tilde{e}_{pr} - \tilde{e}_{pu}) > g$$

Consequently,

$$I_3 < I_m < I_4.$$

Substituting I_4 in $\varphi(I_t)$, we get that $\varphi(I_4) > 0$ if and only if $(1-\psi)(\tilde{e}_{pr} - \tilde{e}_{pu}) > g$. And since we know that for $I_t > \bar{I}_3$ we have $\varphi(I_t) < 0$, we deduce that $I_4 < \bar{I}_3$ □

A.5 Economy 2

In this economy, we have four different groups identified by the threshold I_2, I_3, and I_4. Parents' decisions are made based on the level of their income and the social class they belong to. We summarize the four social in this economy as follows:

- *Poor:* has income below I_2, does not give bequest and invests $e^i_{pu,t}$ in public education.
- *Lower middle class:* has income above I_2 and below I_3, gives bequest $b^i_{pu,t+1}$ and invests e^i_{pu} in public education.
- *Upper middle class:* has income above I_3 and below I_4, does not give bequest and invests $e^i_{pr,t}$ in private education.
- *Rich:* has income above I_4, gives bequest $b^i_{pr,t+1}$ and invests e^i_{pr} in private education.

The marginal effect of government spending in public education g

Differentiating I_2 and I_3, we get:

$$\frac{\partial I_2}{\partial g} = -1 < 0 \tag{2.62}$$

and,

$$\frac{dI_3}{dg} = -\frac{\frac{\partial\varphi}{\partial g}}{\frac{\partial\varphi}{\partial I}} = \frac{(1+\rho+\beta)I}{(1+\rho+\psi\beta)(I-\tilde{e}_{pu}+\frac{g+\tilde{e}_{pu}}{\psi}) - (1+\rho+\beta)I} > 0 \tag{2.63}$$

(2.62) and (2.63) imply that the size of the poor class and the upper middle class reduces and the size of the lower middle class increases as g increases. In contrast, the size of the rich class remains unchanged since I_4 does not depend on government expenditure. The change in policy also impacts the optimal level of investment in public education for both the poor and the lower middle class. This partially explains why the size of the first social class adjusts simultaneously. As I_2 decreases, a marginal group from the poor class finds it optimal to switch to the lower middle class. Similarly, the decrease in the optimal investment in public education $\tilde{e}_{pu}$ instigates a group in the upper middle class to switch to the lower middle class as investing in public education and giving bequest gives more utility now. Note that an increase in government expenditure increases the optimal level of consumption for the poor and the lower middle class. In terms of the income of the next generations, we take into consideration what happens to bequest. Since this latter does not change for both the lower middle class and the rich, I can conclude that the future income of the lower middle class and the future income of the rich remain unchanged. The future income of the lower middle class does not get affected because education and government spending are perfect substitutes, whereas the future income of the rich is unaffected simply because spending on private education and bequests given by this social class do not depend on government spending.[13][14]

The marginal effect of taxation τ

$$\frac{\partial \hat{e}_{pu}}{\partial \tau} = \frac{-1}{(1-\psi)(1-\tau)} \left(\frac{(1-\tau)w\alpha\psi}{R} \right)^{\frac{1}{1-\psi}} \tag{2.64}$$

$$\frac{\partial \hat{e}_{pr}}{\partial \tau} = -\frac{w\psi}{R(1-\psi)} \hat{e}_{pr}^{\psi} \tag{2.65}$$

$$\frac{\partial I_2}{\partial \tau} = -\frac{(1+\rho+\psi\beta)w\alpha}{\beta(1-\psi)} \left(\hat{e}_{pu} + \bar{g}\right)^{\psi} \tag{2.66}$$

$$\frac{dI_3}{d\tau} = -\frac{\frac{\partial \varphi}{\partial \tau}}{\frac{\partial \varphi}{\partial I_3}} = -\left(\frac{\frac{I_3}{\psi(1-\tau)}[(1+\rho+\beta)\tilde{e}_{pu} - \beta\psi(I_3 - \tilde{e}_{pu} + \frac{\bar{g}+\tilde{e}_{pu}}{\psi})]}{(1+\rho+\psi\beta)(I_3 - \tilde{e}_{pu} + \frac{\bar{g}+\tilde{e}_{pu}}{\psi}) - (1+\rho+\beta)I_3} \right) \tag{2.67}$$

The marginal effect of the quality of public education α

$$\frac{\partial I_2}{\partial \alpha} = \frac{1+\rho+\psi\beta}{\psi\beta\alpha(1-\psi)} \left(\tilde{e}_{pu} + g\right) \tag{2.68}$$

$$\frac{dI_3}{d\alpha} = -\frac{\frac{\partial \varphi}{\partial \tau}}{\frac{\partial \varphi}{\partial I_3}} = \frac{\frac{I_3}{\psi\alpha}(1+\rho+\beta)\tilde{e}_{pu}}{(1+\rho+\psi\beta)(I_3 - \tilde{e}_{pu} + \frac{g+\tilde{e}_{pu}}{\psi}) - (1+\rho+\beta)I_3} \tag{2.69}$$

$$\frac{\partial \hat{e}_{pu}}{\partial \alpha} = \frac{1}{(1-\psi)\alpha} \left(\frac{(1-\tau)w\alpha\psi}{R} \right)^{\frac{1}{1-\psi}} \tag{2.70}$$

[13] It should be noted that government spending and investment in public education are imperfect substitutes for social class 1 and perfect substitutes for social class 2. See (2.17), (2.22) and (2.3)

[14] I show the impact of a change in tax and education productivity in the Appendix.

A.6 Data

- All dollar measures are at constant dollars year 2019 using the R-CPI-U-RS produced by the U.S. Bureau of Labor Statistics.
- Data on public elementary-secondary education finance for the years 2005-2019 is collected from the Annual Survey of School System Finances.
- Household measures are collected from American Community Surveys (ACS) from 2005 to 2019.

A.7 Robustness

Table 2.2: Choice of Private Schooling on Total public spending per capita

	Middle class	Rich	Poor	Middle class	Rich	Poor
	(1)	(2)	(3)	(4)	(5)	(6)
Total Pub spending per capita	-0.385***	-0.126	-0.264	-0.353***	-0.094	-0.250
	(-4.798)	(-1.333)	(-1.540)	(-3.948)	(-1.024)	(-1.389)
Household Income	0.690***	0.860***	0.008	0.406***	0.783***	-0.010
	(14.782)	(29.417)	(0.636)	(13.766)	(28.000)	(-1.479)
Household educ	-	-	-	0.679***	0.621***	0.759***
				(23.582)	(15.135)	(19.952)
Metropolitan dummy	0.283***	0.405***	0.176***	0.326***	0.385***	0.335***
	(7.871)	(8.920)	(5.395)	(9.309)	(8.615)	(7.464)
Observations	2633037	1318289	1287456	2633037	1318289	1287456
Time fixed effects	✓	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓	✓
Race controls	✗	✗	✗	✓	✓	✓

Notes: The dependent variable is a dummy that takes 1 if the household has at least one child in private school. The total public expenditure per capita and household income are expressed in logs of constant 2019 dollars. Covariates include household income and race dummies as well as time and age-fixed effects. The poor and rich groups represent the first and last quartiles, respectively. The middle class represents the second and third quartiles. For data sources and summary statistics, see Appendix. Standard errors are clustered at the state level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

3 Inequality and Redistribution: Evidence from the U.S. School Districts

3.1 Introduction

Income inequality has been steadily rising throughout most parts of the world. This trend raises legitimate questions about the capacity to redistribute as well as the mechanisms that are deployed to this end. Such questions become even more important when redistribution is aimed at building productive resources that could potentially mitigate the original widening in the income distribution. In view of the essential role played by initial human capital for building further skills and stable lifelong earnings, public provision of basic education constitutes a critical area to understand. Moreover, it serves as one of the most significant forms of redistribution in society.

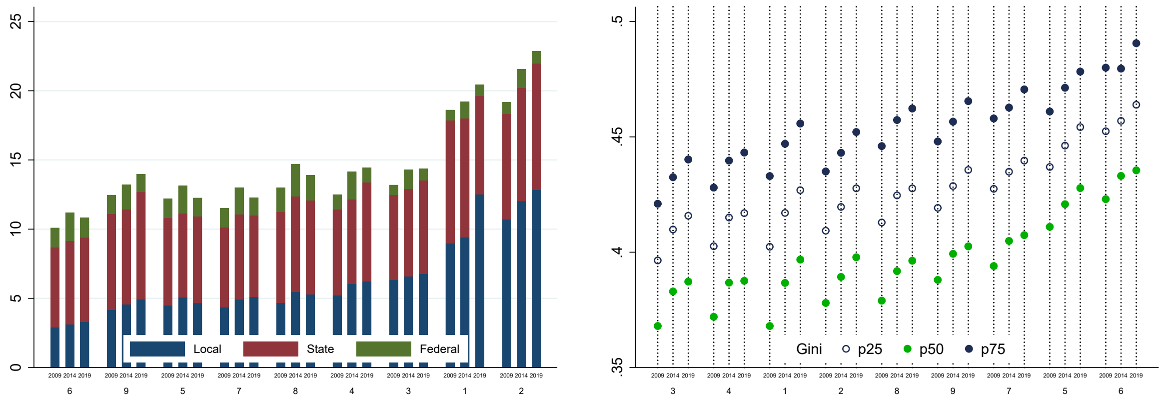
In this paper, we study public education provision across US school districts to shed light on the response of productive redistribution in the face of ever higher income inequality. Panel (a) in Figure 1 shows the evolution of public education funding in the United States by region, separated by source of funding. Public education provision is highly heterogeneous, with local and state spending accounting for large shares. For instance, during the period 2014-2019, the share of local sources[1] in total primary and secondary school funding varies between 0.39% and 97.63% across school districts, with a mean of 43.3%.

On the other hand, income inequality is high relative to other developed economies and rising in all regions between 2009 and 2019 (panel (b) in Figure 1). Together, these features make the universe of American school districts an appropriate setting to identify causal effects of inequality on redistribution and to shed light on the political economy mechanisms that underlie such effects.

Existing theoretical work suggests opposing forces are at work. On the one hand, median voter models pioneered by Meltzer and Richard (1981) highlight the positive effect on inequality on redistribution. On the other hand, seminal political economy models in

[1] The largest portion of local funding for public schools comes from property taxes. Each school district has the authority to levy taxes on properties within its boundaries. Tax rates are set by the local government or school board.

Figure 3.1: Public education funding and income inequality



(a) (b)

Panel (a) Public education funding (in thousands of dollars), by region. Source: Annual Survey of School System Finances. Panel (b) School district Gini coefficients, by region. Source: ACS. US Census regions are from 1 to 9: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain and Pacific.

Bénabou (1996, 2000) suggest more heterogeneous societies may reduce agreement between voters on which public goods to provide and how to fund them. As in many cases private alternatives are within the reach of rich households, their opting out may dovetail with poor households aversion to higher tax burdens, thus giving rise to ”ends against the middle” equilibria described, for example, in Epple and Romano (1996). These general arguments are further complicated in the case of public education provision by quantity-quality trade-offs, first studied by de la Croix and Doepke (2009b) in a framework with endogenous fertility and opting out of public education. Building on the latter, Arcalean and Schiopu (2016) show that keeping the tax base constant, an increase in the spread of the income distribution has non-monotonic effects, depending on the mean income per capita. A mean preserving spread in income increases the mass in the tails of the distribution. Rich families are more likely to opt out of the public education system, thus contributing to a higher spending per pupil in public schools. On the other had, poor households have higher fertility, so an increase in the mass of households of this type decreases the amount of resources available per student. Thus, in rich economies, the opting out effects dominate, so public spending per pupil goes up when income inequality increases. The opposite happens in poor economies, where there is an associated increase in public school enrollment that lowers spending per pupil. Similar non-monotonicities arise in Benzidia et al. (2024) for measures of income polarization/kurtosis.

The empirical literature provides a similarly mixed picture. For example, Lindert (1996) finds a negative relationship between inequality and public education spending in a sample of OECD countries. Using U.S. state level data, de la Croix and Doepke (2009b) find that higher inequality is positively associated with public spending per student and negatively correlated with public spending per capita. Boustan et al. (2013) and Corcoran and Evans (2009) find that rising inequality within U.S. school districts is associated with higher local revenues per pupil.

Identifying causal effects of changes in the income inequality on redistribution is fraught with difficulties. Some are empirical in nature and range from omitted variable bias in country level data to endogenous sorting across smaller units, such as states or districts. Others, not less important, stem from the reduced-form nature of these attempts that exposes them to model misspecification.

A particular source of concern when studying shifts in income distributions is the distinction between changes in its shape from those in its location. Whereas the former can be traced to various inequality statistics, such as the Gini coefficient, the 90/50 or the 10/50 ratios, the latter are typically identified by including the median or mean income in the estimating regression.

Focusing on the effects of inequality on redistribution, most of the existing empirical work, for example Boustan et al. (2013) and Corcoran and Evans (2009), has focused on the median income per capita as a proxy for the preferences of the decisive voter. However, recent evidence on political participation casts doubts on the validity of this choice, at least in the context of U.S. politics, which is relevant to our exercise. Indeed, at both national and local level, political participation indicators such as voter registration and turnout are positively correlated with income but also with socioeconomic status. (Verba et al. (1995), Rosenstone and Hansen (1993), Morlan (1984), Hajnal and Lewis (2003))

From an empirical perspective, these findings imply the voter with the median income is less likely to be decisive. Furthermore, they seem to suggest different theoretical approaches may be needed to describe the complex determination of political equilibria.

Focusing on the political economy of public education provision, probabilistic voting models have been used to shed light on such complexities, from fertility differences and opting out (de la Croix and Doepke (2009b)), residential choice (Melindi-Ghidi (2018)) or indeed political power (Arcalean and Schiopu (2016)). In these models, the political economy equilibrium that emerges depends on the mean income per capita, reflecting the tax base, and on some measure of income dispersion owing to the different types of heterogeneity involved. Within this literature, Arcalean and Schiopu (2016) use a model of endogenous fertility and school choice to show that both tax base and inequality changes may have different effects, depending on the mean income per capita in the economy.

Last but not least, in a distribution that is typically positively skewed, as it is the case with the income distribution, the median income is necessarily correlated with any inequality measure. Thus, any change in inequality will also be reflected in the median income while the mean income tracks only changes in total income, which in our context translates as a change in the tax base.

Based on these empirical and theoretical arguments, we conclude that the relevant statistic for the location of the income distribution is the tax base, proxied by the mean income per capita in the economy.

Our findings confirm that controlling for the tax base of the school district is critical in identifying the effect of inequality on redistribution. Different from most recent literature, we first show that, on average, higher income heterogeneity lowers local revenues per student. Results are robust to a variety of specifications and alternative inequality measures. In particular, we show that the decline of local revenues per pupil in the face of higher income dispersion is mirrored by a similar, albeit smaller decline of total current expenditure per per pupil. This suggests that despite the rising importance of state and federal level outlays, local redistribution matters for the funding of public education.

Providing a comprehensive analysis of the mechanisms that could be behind our results is beyond the purpose of this paper. We note, however, that they are consistent with predictions from the endogenous fertility and school choice models deployed in Arcalean and Schiopu (2016) who build on de la Croix and Doepke (2009b) to demonstrate heterogeneous effects of inequality on redistribution depending on the average income per capita.

In light of these findings, we study empirically how poor and rich districts respond to higher income inequality. We find that the negative overall response in redistribution is driven by the strong decline within poor districts whereas rich districts display statistically insignificant responses. When we redefine the poverty definition to districts in the lowest quartile, the negative coefficient on inequality roughly doubles in size, in absolute value.

In the following we describe our data and methodology. Next, we show the results from the pooled sample, after which we look at heterogeneous effects. Robustness checks and alternative specifications are relegated to the appendix.

3.2 Data and Methodology

We construct our dataset using three different sources. Education finance data is collected from the annual Survey of School System Finances, which provides detailed financial information (including revenue, expenditure, debt, and assets) for public elementary and secondary school systems across the United States. Economic, demographic, and social covariates are obtained from the American Community Survey – Education Tabulation (ACS-ED). Additionally, we use inequality estimates from the National Historical Geographic Information System (NHGIS). The Local Education Agency Identification Numbers (LEAIDs) were employed to merge these three distinct datasets[2].

The constructed dataset includes three cross-sections covering 5-year periods (2005-2009, 2010-2014, 2015-2019), with each period encompassing 11,877, 12,018, and 11,517 school district observations, respectively [3].

[2]The Local Education Agency Identification Numbers (LEAIDs) are compatible with the unit identification codes used by the National Center for Education Statistics (NCESIDs).

[3]More details in Appendix

Using this data, we estimate the following equation in first differences:

$$\Delta L_{it} = \beta_0 + \beta 1 \ \Delta Ineq_{it} + \beta_2 \ \Delta MeanInc_{it} + \beta_3 \ \Delta Z_{it} + \beta_4 \ X_i + \beta_5 \ T_t + \epsilon \tag{3.1}$$

where L_{it} represents school district local funding per pupil, $MeanInc_{it}$ and $Ineq_{it}$ are estimates of the mean household income and respective inequality, Z_{it} is a vector of control variables described below, while X_i and T_t are school district and time-fixed effects, respectively. Throughout, errors are clustered at the school district level.

The vector of control variables Z_{it} includes the state and federal revenues per student in district i at time t.[4] Given the important skewness in income and expenditure data, all these monetary variables are expressed in logs.

Furthermore, school district size, measured by the number of households is included in addition to a set of socio-demographic variables that have shown to shape local public spending decisions: the share of college educated, the share of over-60 (see Poterba (1997), Harris et al. (2001)), the share of non-white as well as a racial diversity index, computed as a Herfindahl index of population shares (see Alesina et al. (1999), Boustan et al. (2013)).

While first differencing the data removes time invariant unobserved heterogeneities in public finance, including district and time fixed effects effectively allow for both national level shocks as well as district specific trends in funding. Finally, given historical differences in initial economic and social conditions across districts, we allow for separate trends in each of the nine Census regions.

We first estimate equation (3.1) using OLS. However, least squares estimates may be biased due to reverse causality stemming from endogenous sorting across districts. If this is the case, the demographic as well as the economic characteristics of the school districts are likely to be endogenous. The same applies to state and federal funding levels which are, by construction, functions of these characteristics (e.g. poverty level). Additional endogeneity concerns arise from the simultaneity in the determination of the local, state, and federal revenues.

To deal with this potentially important issue, we follow a two-step strategy.

First, while our cross-sections cover five year periods due to the limitations of the American Community Survey in small geographies, financial data is available at yearly frequencies. This allows us to mitigate simultaneity concerns by defining our dependent variables in the last year of the five year interval, whereas state and federal revenue numbers are defined as five year averages.

Second, we use an instrumental variable strategy to deal with the possibility of endogenous sorting across district lines. In the spirit of shift-share (Bartik) instruments also used in Boustan et al. (2013), we create counterfactual binned income distributions at the school district (SD) level that keep the population shares in each bin at their initial levels while allowing the representative income in each bin to grow according to national level trends. These counterfactual distributions, therefore, capture only the exogenous compo-

[4] These variables help mitigate other potential biases in the state level policies, such as for example correlations induced by yardstick competition.

nent of income inequality due to broader trends in economic activity that have shifted the national income distribution and which individual districts are too small to influence. We now describe briefly the procedure used to construct synthetic income distributions in 2010 and 2015 based on the 2005 data. Further details are included in the appendix. We start by converting the endpoints of the 2005 SD income bins into percentiles of the national income distribution, obtained from 2005 ACS micro-data from IPUMS-USA.[5] The percentiles are then projected onto the 2010 and 2015 national income distributions, again obtained from ACS micro-data, generating new, synthetic income cutoffs. SD level population shares from 2005 are then assigned to the synthetic 2010 and 2015 income brackets. [6]

The Gini coefficients of these counterfactual distributions are used as instruments for the observed ones. The correlation between the synthetic and the actual series is 0.76.

We start by considering the average effects of inequality on redistribution, specifically local revenues per pupil. Next, we follow theoretical results in Arcalean and Schiopu (2016) to consider the possibility of heterogeneous effects depending on the mean household income. Along the way, we present an extensive batch of robustness exercises, including, among others, different local finance variables, different inequality measures, and expanded sets of control variables.

3.3 Results from the pooled sample

In line with the literature, we first focus on the Gini coefficient as a measure of inequality and study its effects on the local revenues per pupil. Columns (1)-(3) of table 3.1 display least squares estimates. As expected, an increase in the tax base has positive and significant effects on local public finance, with robust coefficients across specifications. Thus, a one percent increase in the mean income generates, on average, an increase of about 0.2 percent in local revenues per pupil. At the same time, higher inequality lowers the growth rate of local revenues per pupil. The negative effect is robust to the inclusion of controls and region specific trends.

Columns (4)-(6) display results from instrumental variable specifications. While both the tax base and the inequality effects maintain their respective signs and significance levels, the coefficient on inequality becomes more negative, moving from -0.27 in column (3) to -0.34 in column (6). Both federal and state level funding have significant effects on local redistribution across specifications, with federal funding acting as a substitute,

[5] The IPUMS-USA database is provided by Ruggles et al. (2010).

[6] For example, suppose that in a given SD, 12% of the households had an income between $10,000 and $20,000 in 2005. Assume that in the 2005 U.S. income distribution, $10,000 and $20,000 correspond to the $7th$ and the $10th$ percentiles, respectively. In the 2010 U.S. distribution, the $7th$ percentile now corresponds to $15,000 while the $10th$ percentile corresponds to $30,000. Thus, in the synthetic 2010 income distribution of the SD, 12% of the households are assumed to have incomes between $15,000 and $30,000.

while state level revenues complement local revenues. Coefficients are stable, including when allowing for heterogeneous evolutions across Census regions.

Since the Gini coefficient tends to be more sensitive to changes in the middle of the distribution, we use the coefficient of variation (CV) and the log of standard deviation to test the robustness of our results to alternative measures that respond more to changes in the tails. Results in tables 3.2 and 3.3 show similarly negative and significant effects on local redistribution in response to higher inequality across both least squares and instrumental variable specifications.

While local revenues per pupil might be lowered by the rise of inequality, total spending per pupil, which is the actual input into the production of human capital could still be unaffected. We therefore re-estimate equation (1) with total current spending per pupil as our dependent variable. Results, displayed in table 3.4 show the coefficient on total expenditure per pupil is still negative and signifcant, albeit smaller in absolute value. Thus, state and federal level redistribution reduce the impact of inequality on per pupil current expenditure, the latter is still affected negatively by an increase in income dispersion, suggesting changes in local public finance variables have real effects on spending.

3.4 Heterogeneous effects

We now turn to explore more in depth the link between local redistribution and inequality at different levels of household income levels. For this, we split the sample according to the mean household income in each period, with districts above (below) the median value being labeled rich and poor respectively.

Table 3.5 contrasts these two subsamples across the same specifications used in table 3.1. While tax base effects remain positive and similarly sized across income groups, the response to inequality changes is markedly different in poor and rich districts. Whereas above median income districts still display negative coefficients, they turn out to be insignificant and lower in absolute value. At the same time, the coefficients in below median income districts remain negative and significant. In the least squares specifications, they have similar magnitudes in absolute value. However, in specifications that correct for endogeneity, coefficients roughly quadruple in absolute value relative to their analogues in the pooled sample.

We perform the same sub-sample analysis using the coefficient of variation and the standard deviation as alternative measures of inequality. Tables 3.7 and 3.8 in Appendix C.2 bring additional evidence on the heterogeneous response of redistribution to the rise in income dispersion. For both measures, it is the lower income districts that drive the negative effect.

Next, we consider a different threshold in our definition of rich and poor districts. In particular, we define as poor districts below the 1st income quartile. Results, reported in tables 3.9-3.11 in appendix C.2 strengthen our finding on the existence of heterogeneous effects. Focusing on changes in the Gini coefficient, districts in the poorest quartile show

Table 3.1: Local Revenues per Pupil and Income Inequality - Gini

	LS	LS	LS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	-0.284***	-0.291***	-0.268***	-0.423**	-0.429**	-0.339*
	(-3.942)	(-3.966)	(-3.680)	(-2.458)	(-2.417)	(-1.911)
Mean Income	0.222***	0.226***	0.192***	0.243***	0.247***	0.204***
	(7.997)	(7.939)	(6.624)	(6.600)	(6.375)	(5.057)
Federal Revenue	-0.024***	-0.024***	-0.027***	-0.024***	-0.024***	-0.027***
	(-4.261)	(-4.341)	(-4.455)	(-4.220)	(-4.306)	(-4.432)
State Revenue	0.049***	0.049***	0.050***	0.049***	0.049***	0.050***
	(5.511)	(5.510)	(5.565)	(5.398)	(5.391)	(5.519)
No. of Households	0.006	0.007	0.006	0.008	0.010	0.008
	(0.189)	(0.238)	(0.196)	(0.265)	(0.319)	(0.244)
Share Over 65		-0.055	3.555		2.399	4.830
		(-0.005)	(0.333)		(0.219)	(0.444)
Share College		-2.285	-0.703		-3.758	-1.466
		(-0.245)	(-0.076)		(-0.392)	(-0.154)
Share Non-White		7.706	-5.720		7.977	-5.494
		(0.461)	(-0.346)		(0.478)	(-0.333)
Diversity Index		0.133	0.173		0.134	0.173
		(1.077)	(1.415)		(1.079)	(1.414)
Observations	21208	21208	21208	21208	21208	21208
Time fixed effects	✓	✓	✓	✓	✓	✓
Region fixed effects	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✓	✗	✗	✓
R-squared	.094	.096	.13	.014	.016	.013
K-P LM statistic	-	-	-	25.5	25.5	25.3
p-value	-	-	-	4.5e-07	4.4e-07	5.0e-07

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Inequality is given by the Gini coefficient at the school district level. Covariates include time and school district fixed effects. Mean income, federal revenue, state revenue, and number of households in each school district are in natural logs. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.2: Local Revenues per Pupil and Income Inequality - CV

	LS	LS	LS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
CV	-0.065***	-0.066***	-0.063***	-0.044**	-0.045**	-0.043**
	(-4.143)	(-4.137)	(-3.946)	(-2.250)	(-2.244)	(-2.160)
Mean Income	0.228***	0.231***	0.199***	0.212***	0.214***	0.183***
	(8.038)	(7.962)	(6.713)	(7.185)	(7.061)	(5.837)
Federal Revenue	-0.024***	-0.025***	-0.027***	-0.024***	-0.025***	-0.027***
	(-4.294)	(-4.374)	(-4.466)	(-4.303)	(-4.381)	(-4.476)
State Revenue	0.050***	0.050***	0.051***	0.050***	0.050***	0.051***
	(5.577)	(5.580)	(5.643)	(5.612)	(5.618)	(5.658)
No. of Households	0.006	0.007	0.006	0.004	0.005	0.004
	(0.178)	(0.224)	(0.189)	(0.131)	(0.172)	(0.134)
Share Over 65	-	-0.262	3.510	-	-1.878	2.004
		(-0.024)	(0.329)		(-0.176)	(0.189)
Share College	-	-2.859	-1.298	-	-1.659	-0.201
		(-0.306)	(-0.140)		(-0.177)	(-0.022)
Share Non-White	-	6.676	-6.569	-	6.828	-6.571
		(0.401)	(-0.400)		(0.409)	(-0.399)
Diveristy Index	-	0.139	0.179	-	0.137	0.177
		(1.133)	(1.466)		(1.111)	(1.452)
Observations	21208	21208	21208	21208	21208	21208
Time fixed effects	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✓	✗	✗	✓
R-squared	.095	.096	.13	.015	.017	.014
K-P LM statistic	-	-	-	379	377	376
p-value	-	-	-	2.2e-84	5.7e-84	1.0e-83

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Inequality is measured as the coefficient of variation (CV) of the school district income distribution. Covariates include time and school district fixed effects. Mean income, federal revenue, state revenue, and number of households in each school district are in natural logs. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.3: Local Revenues per Pupil and Income Inequality - Ln(st.dev.)

	LS	LS	LS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Ln SD	-0.063***	-0.064***	-0.060***	-0.040**	-0.041**	-0.041**
	(-4.185)	(-4.212)	(-4.008)	(-1.987)	(-1.974)	(-1.978)
Mean Income	0.293***	0.299***	0.262***	0.253***	0.256***	0.226***
	(7.701)	(7.637)	(6.624)	(5.563)	(5.426)	(4.657)
Federal Revenue	-0.024***	-0.025***	-0.027***	-0.024***	-0.025***	-0.027***
	(-4.296)	(-4.378)	(-4.470)	(-4.305)	(-4.384)	(-4.479)
State Revenue	0.050***	0.050***	0.051***	0.050***	0.050***	0.051***
	(5.542)	(5.542)	(5.603)	(5.587)	(5.592)	(5.629)
No. of Households	0.006	0.008	0.006	0.004	0.006	0.004
	(0.189)	(0.238)	(0.200)	(0.134)	(0.175)	(0.140)
Share Over 65	-	0.357	4.047	-	-1.662	2.348
		(0.033)	(0.380)		(-0.155)	(0.221)
Share College	-	-3.031	-1.443	-	-1.638	-0.283
		(-0.325)	(-0.156)		(-0.174)	(-0.030)
Share Non-White	-	7.178	-6.119	-	7.164	-6.265
		(0.431)	(-0.372)		(0.429)	(-0.380)
Diversity Index	-	0.137	0.177	-	0.135	0.176
		(1.113)	(1.448)		(1.096)	(1.440)
Observations	21208	21208	21208	21208	21208	21208
Time fixed effects	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✓	✗	✗	✓
R-squared	.095	.096	.13	.015	.017	.014
K-P LM statistic	-	-	-	843	846	842
p-value	-	-	-	2.8e-185	4.9e-186	3.4e-185

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Inequality is measured as the standard deviation of the school district income distribution, in constant 2019 dollars, in logs. Covariates include time and school district fixed effects. Mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.4: Total Current Expenditure per Pupil and Income Inequality - Gini

	LS	LS	LS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	-0.094***	-0.101***	-0.037	-0.165**	-0.168**	-0.121*
	(-3.145)	(-3.359)	(-1.273)	(-2.264)	(-2.237)	(-1.713)
Mean Income	0.134***	0.138***	0.068***	0.144***	0.147***	0.081***
	(11.915)	(12.010)	(6.135)	(9.924)	(9.720)	(5.345)
Federal Revenue	0.017***	0.016***	-0.006***	0.017***	0.016***	-0.006**
	(6.902)	(6.690)	(-2.636)	(6.950)	(6.731)	(-2.574)
State Revenue	0.033***	0.032***	0.026***	0.032***	0.032***	0.025***
	(8.058)	(8.019)	(6.901)	(7.977)	(7.938)	(6.843)
No. of Households	-0.002	0.002	0.001	-0.001	0.003	0.003
	(-0.155)	(0.151)	(0.074)	(-0.062)	(0.244)	(0.213)
Share Over 65	-	0.075	0.067	-	0.087*	0.083*
		(1.586)	(1.492)		(1.767)	(1.769)
Share College	-	0.001	0.005	-	-0.006	-0.003
		(0.016)	(0.144)		(-0.149)	(-0.080)
Share Non-White	-	0.462***	0.203***	-	0.462***	0.205***
		(6.421)	(3.505)		(6.423)	(3.524)
Diveristy Index	-	-0.272***	-0.153***	-	-0.271***	-0.152***
		(-5.261)	(-3.524)		(-5.243)	(-3.511)
Observations	21244	21244	21244	21244	21244	21244
Time fixed effects	✓	✓	✓	✓	✓	✓
Region fixed effects	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✓	✗	✗	✓
R-squared	.32	.32	.43	.032	.039	.013
K-P LM statistic	-	-	-	667	656	668
p-value	-	-	-	4.1e-147	1.1e-144	3.3e-147

Notes: The dependent variable is the change in total current expenditure per student, expressed in logs of constant 2019 dollars. Inequality is given by the Gini coefficient at school district level. Covariates include time and school district fixed effects. Mean income, federal revenue, state revenue, and number of households in each school district are in natural logs. For data sources and summary statistics see Appendix A. Standard errors are clustered at school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.5: Local Revenues per Pupil and Income Inequality (Gini) - Heterogeneous Effects

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gini	-0.244***	-0.818**	-0.109	-0.296	-0.211**	-0.825**	-0.058	-0.199
	(-2.918)	(-2.454)	(-1.285)	(-1.394)	(-2.527)	(-2.418)	(-0.680)	(-0.934)
Mean Income	0.249***	0.256***	0.256***	0.280***	0.226***	0.242***	0.215***	0.246***
	(7.277)	(4.129)	(6.729)	(4.908)	(6.487)	(3.721)	(5.654)	(4.190)
No. of Households	-0.060**	0.029	-0.041	-0.019	-0.067**	0.025	-0.053	-0.012
	(-1.982)	(0.603)	(-1.328)	(-0.381)	(-2.180)	(0.507)	(-1.640)	(-0.232)
Federal Revenue	-0.021**	0.011	-0.041***	-0.050***	-0.019*	0.009	-0.042***	-0.046***
	(-2.370)	(1.004)	(-6.272)	(-7.175)	(-1.953)	(0.749)	(-6.072)	(-6.347)
State Revenue	0.009	0.054***	0.011	0.041***	-0.001	0.047**	0.015*	0.046***
	(0.605)	(2.781)	(1.302)	(3.728)	(-0.074)	(2.327)	(1.765)	(4.146)
Share Over 65	33.348***	15.953	8.799	-18.486	38.986***	19.996	5.852	-14.697
	(3.256)	(0.955)	(0.745)	(-1.026)	(3.823)	(1.197)	(0.498)	(-0.825)
Share College	-1.903	-6.495	-5.899	2.876	-1.042	-7.684	-1.173	6.292
	(-0.157)	(-0.372)	(-0.637)	(0.233)	(-0.086)	(-0.440)	(-0.127)	(0.519)
Share Non-White	-12.817	-26.086	52.278***	81.205***	-13.967	-26.393	18.145	36.469
	(-0.876)	(-1.215)	(3.328)	(3.368)	(-0.957)	(-1.225)	(1.171)	(1.521)
Diveristy Index	0.204*	0.420**	-0.292**	-0.409**	0.208*	0.381**	-0.090	-0.188
	(1.803)	(2.486)	(-2.463)	(-2.299)	(1.832)	(2.245)	(-0.761)	(-1.067)
Observations	10972	8830	11278	9160	10972	8830	11278	9160
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
R-squared	.045	.01	.065	.033	.063	.0069	.097	.027
K-P LM statistic	-	201	-	358	-	205	-	358
p-value	-	1.3e-45	-	7.0e-80	-	2.1e-46	-	6.6e-80

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the median value every year. Inequality is given by the Gini coefficient at the school district level. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

a negative and significant effect which is twice as large as the effect among the below median districts. Again, this result holds across specifications and all our three inequality measures.

Taken together, results so far lead us to conclude that the overall negative response to higher inequality is the result of muted responses in rich districts that become negative and significant as average income per capita declines.

Our results are in stark contrast with previous literature estimating the effect of changes in Gini coefficients controlling for the median income per capita. We therefore re-estimate equation (1) adding the median income per capita as a control. Tables 3.12 - 3.17 show that irrespective of the measure used, the sign and magnitude of the inequality coefficients do not change while the median has an insignificant coefficient conditional on the mean income per capita.

3.5 Conclusion

In this paper we take a fresh look at the question of inequality and redistribution. Using the case of U.S. school districts over the period 2005-2019, we make three contributions. First, we collect empirical and theoretical arguments from recent literature to argue that the mean income per capita should be included in the estimating equation in order to properly identify the effect of inequality. Second, we show that this effect is negative and significant, in other words redistribution is reduced in the face of higher income dispersion. Not only do local revenues per pupil decline, but, despite the significant and increasing weight of state and federal funding, total current expenditure per pupil is lowered too, pointing at real effects on the provision of public education services. Adding the median income per capita in the school district to the estimation fails to change the sign of the inequality coefficient, while the median income itself remains insignificant conditional on the tax base being controlled for.

Considering the possibility of heterogeneous effects, we further show the average effect is driven by the strong negative response in redistribution in poor school districts, defined as those with an average income per capita below the median value and muted, insignificant responses in richer school districts. Redefining the poverty threshold to include just districts in the lowest quartile, the negative coefficient on inequality doubles in size in absolute value, which points to a gradual worsening of local redistribution as the income per capita declines.

We emphasize that our causal estimates are consistent with the non-monotonic effects demonstrated in recent theoretical work highlighting opting out and differential fertility choices across income groups. Thus, our research has implications for the design of equitable education funding policies, against the background of persistent achievement gaps by socioeconomic status. On a more general note, our results invite a closer look at other redistribution policies deployed in the face of higher income inequality.

Appendix B

B.1 Data

Table 3.6: Summary Statistics

	mean	median	sd	min	max
Mean income by school district	78595.61	69891.49	33063.46	26079.00	466343.00
Median income by school district	62052.40	56053.13	24560.48	14438.00	296177.20
Local Rev. / Student	6634.68	5132.84	5294.77	0	121717.60
Federal Rev. / Student	1301.97	1024.40	1520.95	2.77	105736.36
State Rev. / Student	6823.79	6296.87	3529.56	90.15	94259.42
Share of Public Ed. Enroll.	0.90	0.91	0.07	0.04	1
Diversity Index	0.21	0.15	0.17	0	0.77
Gini Index	0.42	0.42	0.05	0.26	0.71
N. Households	10304	3517	43215	42	3167034
Share over 65	0.16	0.16	0.05	0	0.71
Share with college	0.15	0.13	0.08	0	0.51
Share non-white	0.15	0.08	0.17	0	1
Observations	35412				

Public elementary-secondary education finance

- Data on public elementary-secondary education finance for the years 2005-2019 is collected from the Annual Survey of School System Finances.
- The variables of interest in this data set are :
  - NCESID: NCES Identification Number
  - NAME: School System Name
  - YRDATA: Year of Data
  - ENROLL: Fall Membership (numbers of students)
  - TFEDREV: Total Revenue from Federal Sources
  - TSTREV: Total Revenue from State Sources
  - TLOCREV: Total Revenue from Local Sources
  - TCURELSC: Total current spending for elementary-secondary programs.

  - PPCSTOT: per pupil - Total current spending (Elementary Secondary). Note that this variable is not available for all years. We compute it when missing, following this equation $PPCSTOT = (TCURELSC/ENROLL) * 1000$

- Variables are aggregated by 5-year periods (2005-2009, 2010-2014, and 2015-2019)

American Community Survey – Education Tabulation (ACS-ED)

- Geographic coverage: school district.
- Three 5-year periods 2005-2009, 2010-2015, and 2016-2019.
- Provides economic, demographic, and social data.
  - The economic variables are mean/median income, and The number of households by 10 income bracket (less than $10,000, $10,000 to $14,999, $15,000 to $24,999, $25,000 to $34,999, $35,000 to $49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 to $149,999, $150,000 to $199,999, and $200,000 or more)
  - The demographic variables used are the share of people older than 65 years, the share of nonwhite, and the Diversity index. These variables are computed using the SEX & Age, and Race estimates available in the Total Population universe.
  - The social variable used is share with college in population 25 and over. This variable is computed using education attainment estimates in the Total Population universe. Additionally, public enrollment share at the school district level is calculated using total and public enrollment from the children universe.

National Historical Geographic Information System (NHGIS)

- Geographic coverage: school district.
- three 5 years periods 2005-2009, 2010-2015, and 2016-2019.
- Provides economic data.
  - The economic variables are median income, aggregate household income, household income quintile upper limits/Lower Limit of Top 5 Percent, mean household income of quintiles/mean income of top 5 Percent, shares of aggregate household income by quintile/top 5 Percent Shares of Aggregate Household Income, Gini index of income inequality, and the number of households by 16 income brackets (Less than $10,000, $10,000 to $14,999, $15,000 to $19,999, $20,000 to $24,999, $25,000 to $29,999, $30,000 to $34,999, $35,000 to $39,999, $40,000 to $44,999, $45,000 to $49,999, $50,000 to $59,999, $60,000 to $74,999, $75,000 to $99,999, $100,000 to $124,999, $125,000 to $149,999, $150,000 to $199,999, $200,000 or more). The mean household income by school district is

computed using the total number of households and the aggregate household income in the district.

This dataset enables computing more accurate inequality estimates due to its increased number of bins. The availability of additional variables allows for the creation of finer bins, ultimately leading to more precise synthetic inequality measures.

Consumer Price Index for All Urban Consumers Retrospective Series

All dollar measures are inflated using the R-CPI-U-RS produced by the U.S. Bureau of Labor Statistics.

B.2 Construction of synthetic inequality instrument

The synthetic inequality instrument is constructed following these steps:

- The endpoints of each of the income bins, describing the income distribution at the school district level in the period 2005-2009, are transformed into the percentiles of the aggregate US household income distribution in 2010-2014, and 2015-2019.

- For each district the population counts in bins from the 2005-2009 ACS are stored. The midpoint of each bin is assumed as the representative income for that bin. The top (open-ended) bin is assigned an income consistent with the grand mean. It should be noted that depending on the data set used to compute the synthetic controls, different strategies are applied to assign the "midpoint" of the top open-ended bin. When using the ACS-ED, the bin is assigned an income consistent with the grand mean. To this end, the aggregate household income for the last bin is estimated and divided by the bin‘s population count. In a few cases, this output is below 200,000, so it is replaced with 1.4 * 200,000 (see Autor et al. 2008, Lemieux 2006, von Hippel et al. 2017). If the population is zero, the midpoint is set as the lower bound of the bin. When using NHGIS, the additional data points (namely the lower limit and the mean Household Income of the top 5 percent) are utilized to enhance the granularity and accuracy of the last bin midpoint. Mainly to compute the shape parameter of the Pareto distribution α. subsequently, if the lower limit of the top bin (let's call it x_m) is greater than 200,000, the midpoint is the median $x = x_m * 2^{(1/alpha)}$. If the scale parameter x_m of the Pareto distribution is less than \$200,000, then to find the midpoint of the top bin, we compute the cumulative probability associated with the value \$200,000 by solving the equation $200,000 = x_m \cdot (1 - F(200,000))^{-\frac{1}{\alpha}}$. Once $F(200,000)$ is determined, the cumulative probability associated with the midpoint of the top bin is calculated by $F(x) = \frac{1+F(200,000)}{2}$. Finally, we calculate the value in the distribution corresponding to the cumulative probability $F(x)$ as follows $x = x_m \cdot (1 - F(x))^{-\frac{1}{\alpha}}$.

- The median income for each bin in the 2005-2009 national income distribution is found, and the maximum observed value is used to pin down the median in the top bin. The percentiles corresponding to bin ends and the computed bins‘ medians are stored.

- The percentiles are then mapped to their dollar equivalents on the US household income distribution for the 5 years periods (2010-2014, and 2015-2019) separately.

- New median income levels between the "equivalent" cutoffs are computed. The empirical maximum is used to pin down the median in the top bin.

- Growth rates in these median incomes are computed relative to 2005-2009 levels.

- Using district midpoint incomes in 2005-2009 and bin-specific growth rates, synthetic incomes for 2010-2014 and 2015-2019 are calculated. Finally, Ginis and other inequality statistics (r9050, r1050, median, mean, and sd) are derived using the **ineqdeco** package.

B.3 Robustness

Table 3.7: Local Revenues per Pupil and Income Inequality (CV) - Rich and Poor (above/below median income)

	LS Poor (1)	IV Poor (2)	LS Rich (3)	IV Rich (4)	LS Poor (5)	IV Poor (6)	LS Rich (7)	IV Rich (8)
CV	-0.061***	-0.069***	-0.012	-0.043	-0.053***	-0.067**	0.000	-0.042
	(-4.004)	(-2.594)	(-0.581)	(-1.180)	(-3.544)	(-2.508)	(0.006)	(-1.156)
Mean Income	0.266***	0.208***	0.245***	0.255***	0.242***	0.191***	0.205***	0.237***
	(7.466)	(4.257)	(6.498)	(5.254)	(6.663)	(3.762)	(5.410)	(4.768)
No. of Households	-0.059*	0.023	-0.042	-0.023	-0.066**	0.018	-0.053*	-0.014
	(-1.956)	(0.479)	(-1.361)	(-0.456)	(-2.151)	(0.366)	(-1.665)	(-0.283)
Observations	10972	8830	11278	9160	10972	8830	11278	9160
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.046	.016	.065	.032	.064	.013	.097	.027
K-P LM statistic	-	204	-	238	-	207	-	242
p-value	-	2.4e-46	-	1.1e-53	-	7.4e-47	-	1.7e-54

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the median value every year. Inequality is measured as the coefficient of variation (CV) of the school district income distribution. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.8: Local Revenues per Pupil and Income Inequality (ln(st.dev.)) - Rich and Poor (above/below median income)

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln SD	-0.063***	-0.073**	-0.013	-0.024	-0.056***	-0.071**	-0.001	-0.028
	(-4.019)	(-2.378)	(-0.684)	(-0.769)	(-3.573)	(-2.272)	(-0.069)	(-0.877)
Mean Income	0.334***	0.289***	0.260***	0.272***	0.301***	0.267***	0.208***	0.260***
	(7.378)	(3.923)	(5.294)	(3.910)	(6.581)	(3.523)	(4.190)	(3.640)
No. of Households	-0.059*	0.023	-0.041	-0.023	-0.066**	0.017	-0.053*	-0.014
	(-1.959)	(0.475)	(-1.356)	(-0.455)	(-2.158)	(0.359)	(-1.664)	(-0.278)
Observations	10972	8830	11278	9160	10972	8830	11278	9160
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.046	.016	.065	.032	.064	.013	.097	.026
K-P LM statistic	-	377	-	398	-	380	-	406
p-value	-	5.6e-84	-	1.6e-88	-	1.6e-84	-	2.5e-90

Notes: The table shows estimates The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the median value every year. Inequality is measured as the standard deviation of the school district income distribution, in constant 2019 dollars, in logs. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.9: Local Revenues per Pupil and Income Inequality (Gini)- Alternative Income Threshold (Above/Below First Quartile)

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gini	-0.393***	-1.466**	-0.113*	-0.275	-0.375***	-1.609**	-0.064	-0.155
	(-3.395)	(-2.296)	(-1.650)	(-1.506)	(-3.277)	(-2.425)	(-0.934)	(-0.852)
Mean Income	0.154***	0.289***	0.292***	0.276***	0.153***	0.324***	0.248***	0.225***
	(3.612)	(2.968)	(9.516)	(6.094)	(3.549)	(3.129)	(8.040)	(4.800)
No. of Households	-0.045	0.022	-0.052**	-0.010	-0.044	0.045	-0.065***	-0.020
	(-1.048)	(0.261)	(-2.136)	(-0.291)	(-1.014)	(0.529)	(-2.585)	(-0.549)
Observations	5406	3678	16844	14582	5406	3678	16844	14582
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.045	.0053	.06	.027	.069	-.00014	.085	.022
K-P LM statistic	-	86.7	-	486	-	89.5	-	484
p-value	-	1.3e-20	-	8.6e-108	-	3.1e-21	-	2.8e-107

Notes: The table shows estimates The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the first quartile every year. Inequality is given by the Gini coefficient at the school district level. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.10: Local Revenues per Pupil and Income Inequality (CV)-Alternative Income Threshold (Above/Below First Quartile)

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CV	-0.074***	-0.115***	-0.026*	-0.023	-0.071***	-0.123***	-0.015	-0.018
	(-3.619)	(-2.715)	(-1.841)	(-0.939)	(-3.476)	(-2.845)	(-1.047)	(-0.713)
Mean Income	0.167***	0.228***	0.293***	0.246***	0.165***	0.250***	0.249***	0.211***
	(3.744)	(3.085)	(9.501)	(6.683)	(3.659)	(3.285)	(8.022)	(5.560)
No. of Households	-0.044	0.023	-0.053**	-0.015	-0.043	0.045	-0.066***	-0.022
	(-1.015)	(0.265)	(-2.177)	(-0.412)	(-0.980)	(0.532)	(-2.613)	(-0.621)
Observations	5406	3678	16844	14582	5406	3678	16844	14582
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.045	.028	.06	.026	.069	.028	.085	.021
K-P LM statistic	-	123	-	198	-	127	-	195
p-value	-	1.3e-28	-	5.8e-45	-	1.7e-29	-	3.1e-44

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the first quartile every year. Inequality is measured as the standard deviation of the school district income distribution, in constant 2019 dollars, in logs. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.11: Local Revenues per Pupil and Income Inequality (ln (st.dev.))-Alternative Income Threshold (Above/Below First Quartile)

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln SD	-0.086***	-0.122**	-0.024*	-0.020	-0.081***	-0.133***	-0.013	-0.017
	(-4.003)	(-2.487)	(-1.721)	(-0.832)	(-3.871)	(-2.638)	(-0.929)	(-0.722)
Mean Income	0.261***	0.352***	0.317***	0.267***	0.254***	0.387***	0.261***	0.230***
	(4.620)	(3.108)	(7.951)	(4.854)	(4.439)	(3.288)	(6.496)	(4.093)
No. of Households	-0.044	0.019	-0.053**	-0.014	-0.043	0.041	-0.066***	-0.022
	(-1.025)	(0.221)	(-2.168)	(-0.403)	(-0.991)	(0.483)	(-2.609)	(-0.608)
Observations	5406	3678	16844	14582	5406	3678	16844	14582
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.046	.029	.06	.026	.07	.029	.085	.021
K-P LM statistic	-	212	-	517	-	220	-	511
p-value	-	5.9e-48	-	2.1e-114	-	9.2e-50	-	4.3e-113

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Inequality is measured as the standard deviation of the school district income distribution, in constant 2019 dollars, in logs. Districts are split into Rich and Poor according to the mean income per capita falling above or below the first quartile every year. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.12: Local Revenues per Pupil and Income Inequality (Gini)- with median income

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gini	-0.352*	-1.337*	-0.416**	-0.615	-0.336*	-1.364**	-0.381**	-0.411
	(-1.879)	(-1.954)	(-2.359)	(-1.031)	(-1.806)	(-1.991)	(-2.194)	(-0.732)
Mean Income	0.182**	0.519**	0.373***	0.447	0.162*	0.515**	0.346***	0.357
	(2.169)	(2.011)	(3.935)	(1.483)	(1.930)	(1.964)	(3.715)	(1.406)
Median Income	0.021	-0.273	-0.111	-0.174	0.022	-0.284	-0.106	-0.115
	(0.303)	(-1.182)	(-1.371)	(-0.832)	(0.316)	(-1.243)	(-1.346)	(-0.551)
Observations	8830	8830	9160	9160	8830	8830	9160	9160
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.08	.0047	.13	.033	.1	.0007	.18	.027
K-P LM statistic	-	20.4	-	18.3	-	20.2	-	18.5
p-value	-	6.2e-06	-	.000019	-	6.9e-06	-	.000017

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the median value every year. Inequality is measured as the Gini coefficient of the school district income distribution. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.13: Local Revenues per Pupil and Income Inequality (CV)- with median income

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CD	-0.073***	-0.062*	-0.076*	-0.053	-0.070**	-0.061*	-0.073*	-0.054
	(-2.579)	(-1.948)	(-1.949)	(-0.853)	(-2.476)	(-1.944)	(-1.883)	(-0.965)
Mean Income	0.194**	0.173*	0.313***	0.284*	0.173**	0.158	0.295***	0.271**
	(2.536)	(1.832)	(3.879)	(1.748)	(2.260)	(1.605)	(3.706)	(2.296)
Median Income	0.025	0.041	-0.061	-0.036	0.027	0.038	-0.064	-0.044
	(0.447)	(0.487)	(-0.869)	(-0.374)	(0.473)	(0.470)	(-0.936)	(-0.455)
Observations	8830	8830	9160	9160	8830	8830	9160	9160
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.081	.016	.13	.033	.1	.013	.18	.027
K-P LM statistic	-	20.4	-	22.3	-	20.4	-	22.3
p-value	-	6.3e-06	-	2.4e-06	-	6.4e-06	-	2.4e-06

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the median value every year. Inequality is measured as the coefficient of variation (CV) of the school district income distribution. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.14: Local Revenues per Pupil and Income Inequality (ln (st.dev.)) -with median income

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln SD	-0.083***	-0.067	-0.066*	-0.029	-0.078***	-0.065*	-0.067**	-0.037
	(-2.726)	(-1.626)	(-1.939)	(-0.508)	(-2.607)	(-1.647)	(-1.965)	(-0.740)
Mean Income	0.300***	0.255*	0.383***	0.291	0.272***	0.235	0.370***	0.297*
	(2.895)	(1.766)	(3.456)	(1.276)	(2.631)	(1.626)	(3.381)	(1.712)
Median Income	0.009	0.031	-0.064	-0.017	0.012	0.030	-0.071	-0.034
	(0.152)	(0.334)	(-0.911)	(-0.173)	(0.202)	(0.329)	(-1.039)	(-0.338)
Observations	8830	8830	9160	9160	8830	8830	9160	9160
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.081	.016	.13	.032	.1	.013	.18	.027
K-P LM statistic	-	21.6	-	21.9	-	21.6	-	21.9
p-value	-	3.3e-06	-	2.9e-06	-	3.4e-06	-	2.8e-06

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the median value every year. Inequality is measured as the standard deviation of the school district income distribution, in constant 2019 dollars, in logs. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.15: Local Revenues per Pupil and Income Inequality (Gini)- with median income and alternative income threshold

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gini	-0.559*	-2.568*	-0.245*	-0.376	-0.626**	-2.823**	-0.626**	-0.114
	(-1.935)	(-1.872)	(-1.874)	(-0.891)	(-2.156)	(-2.004)	(-2.156)	(-0.273)
Mean Income	0.208	0.852*	0.279***	0.328**	0.240*	0.945**	0.240*	0.204
	(1.639)	(1.857)	(4.148)	(1.993)	(1.879)	(2.013)	(1.879)	(1.241)
Median Income	-0.019	-0.600	-0.012	-0.053	-0.033	-0.666	-0.033	0.022
	(-0.179)	(-1.472)	(-0.208)	(-0.383)	(-0.315)	(-1.608)	(-0.315)	(0.157)
Observations	3678	3678	14582	14582	3678	3678	3678	14582
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.087	-.014	.11	.027	.12	-.023	.12	.022
K-P LM statistic	-	82.8	-	284	-	82.2	-	283
p-value	-	9.1e-20	-	1.2e-63	-	1.2e-19	-	1.5e-63

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the first quartile every year. Inequality is measured as the Gini coefficient of the school district income distribution. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.16: Local Revenues per Pupil and Income Inequality (CV)- with median income and alternative income threshold

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CD	-0.106***	-0.118**	-0.044*	-0.012	-0.114***	-0.127**	-0.114***	-0.007
	(-2.615)	(-2.210)	(-1.690)	(-0.376)	(-2.783)	(-2.357)	(-2.783)	(-0.210)
Mean Income	0.225**	0.246*	0.251***	0.206***	0.251**	0.274**	0.251**	0.171***
	(2.060)	(1.807)	(4.382)	(3.223)	(2.285)	(2.016)	(2.285)	(2.676)
Median Income	-0.005	-0.022	0.013	0.050	-0.010	-0.027	-0.010	0.049
	(-0.064)	(-0.207)	(0.264)	(0.910)	(-0.124)	(-0.269)	(-0.124)	(0.907)
Observations	3678	3678	14582	14582	3678	3678	3678	14582
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.09	.028	.11	.026	.12	.028	.12	.021
K-P LM statistic	-	137	-	169	-	138	-	167
p-value	-	1.5e-31	-	1.3e-38	-	9.3e-32	-	3.4e-38

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the first quartile every year. Inequality is measured as the coefficient of variation (CV) of the school district income distribution. Covariates include time and school district fixed effects. mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at the school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

Table 3.17: Local Revenues per Pupil and Income Inequality - (ln(st.dev.))- with median income and alternative income threshold

	LS Poor	IV Poor	LS Rich	IV Rich	LS Poor	IV Poor	LS Rich	IV Rich
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln SD	-0.131***	-0.128**	-0.040*	-0.006	-0.139***	-0.141**	-0.139***	-0.005
	(-2.986)	(-1.982)	(-1.701)	(-0.177)	(-3.166)	(-2.153)	(-3.166)	(-0.140)
Mean Income	0.397***	0.386*	0.296***	0.204**	0.430***	0.433**	0.430***	0.174*
	(2.713)	(1.853)	(3.767)	(2.042)	(2.906)	(2.065)	(2.906)	(1.733)
Median Income	-0.038	-0.033	0.010	0.056	-0.043	-0.044	-0.043	0.051
	(-0.437)	(-0.285)	(0.202)	(0.948)	(-0.495)	(-0.391)	(-0.495)	(0.863)
Observations	3678	3678	14582	14582	3678	3678	3678	14582
Time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
SD fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional trends	✗	✗	✗	✗	✓	✓	✓	✓
Funding controls	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	.092	.029	.11	.026	.13	.029	.13	.021
K-P LM statistic	-	197	-	460	-	199	-	456
p-value	-	8.3e-45	-	4.3e-102	-	3.6e-45	-	2.9e-101

Notes: The dependent variable is the change in local public revenues per student, expressed in logs of constant 2019 dollars. Districts are split into Rich and Poor according to the mean income per capita falling above or below the first quartile every year. Inequality is measured as the standard deviation of the school district income distribution, in constant 2019 dollars, in logs. Covariates include time and school district fixed effects. Mean income, federal revenue, state revenue, and number of households in each school district are in natural log. For data sources and summary statistics see Appendix A. Standard errors are clustered at school district level and are reported in brackets. * indicates significance at the 10 percent level, ** indicates significance at the 5 percent level, *** indicates significance at the 1 percent level.

4 Marry for Money or Time? Explaining New Marriage Trends in the U.S.

4.1 Introduction

Over the past several decades, aggregate marriage rates have been declining worldwide, a trend with far-reaching social and economic implications for household formations, fertility rates, labor market dynamics, and inequality. Understanding the factors driving changes in marriage patterns is crucial for policymakers and researchers alike. Recently, Moro et al. (2017), using cross-country panel data on marital statistics from 1900 to 2000 for 16 OECD countries, document that the fraction of married individuals follows a hump-shaped pattern. The peak in marriage rates, occurring between 1960 and 1980, underscores a turning point after which marriage rates began their steady decline. Many other economists highlighted the same trend.

It has been understood now that the decline in marriage has been accompanied by several stylized facts. First, the literature emphasizes that marriage patterns are not homogeneous across groups, particularly among different education levels (Chiappori et al., 2018; Chiappori et al.,2020) and racial groups (Caucutt et al., 2021). In the United States for example, the decline in marriage rates has been most apparent among non-college-educated and Black Americans (Greenwood et al.,2016; Caucutt et al.,2021). Another established stylized fact is the rise in assortative mating, particularly an increase in the fraction of marriages in which both partners share the same educational level (Greenwood et al.,2016).

The two main underlying forces considered by the literature are technological progress in household production and changes in the labor market that began in the 1960s, such as educational attainment and female labor force participation. (Stevenson and Wolfers, 2007; Greenwood and Guner, 2009; Greenwood et al., 2014; Greenwood et al., 2016). Some other drivers include technological improvement in contraceptives and changed social perception towards sex outside wedlock (Greenwood and Guner, 2010).

The purpose of this paper is to investigate the impact of two specific economic factors on recent marriage patterns in the United States; factors that have not been extensively explored by the existing literature, specifically the effects of changes in the wage structure and work hours on household formation. To this end, I utilize data from the U.S. Census and the American Community Survey (ACS) from 2000 to 2021, including the 1% ACS and 5% Census samples provided through IPUMS-USA (Ruggles et al., 2024). I divide my data into two income groups for each gender, using the median income specific to each gender as the threshold to separate the groups. For example, men are divided into those falling below and those above the median income for men, and the same approach is applied to women. This method ensures that individuals are compared to their gender. I then compute the median wage, the average hours worked, and the marriage fraction for each group.

This paper offers two main contributions. First, it presents new stylized facts about marriage patterns in the United States, highlighting how different income groups within each gender experienced divergent marriage trends. Second, it develops a general equilibrium model of marriage that incorporates exogenous changes in the wage structure and working hours. This model allows us to assess the extent to which these economic factors influence the observed new marriage patterns.

I document a new stylized fact on marriage patterns and relate it to three other patterns established by previous literature. First, I show a reduction in the gender wage gap and an increase in wage inequality within gender. Second, I observe a decrease in men's work hours accompanied by an increase in women's work hours. Third, I document an increase in couples in which both partners are high earners, indicative of positive assortative mating in income. Fourth, I show that the income groups do not all experience the same decline in marriage rates. Notably, and contrary to existing literature, while the fractions of married individuals for all income groups of men and for low-income women are declining, the fraction of high-income women who are married is increasing. This finding is a novelty that becomes even more apparent when I divide the population into more income groups.

I then build a model of marriage featuring agents with different wage and hour levels. I calibrate the model to replicate the observed marriage fraction and marital sorting in the baseline year 2000. Subsequently, I test the model's performance in accounting for changes in the marriage rates of the different groups in the year 2021, explained by two channels: changes in wages and hours worked. Then, I run a decomposition exercise to assess the two channels. I find that although changes in the wage structure contribute to the declining marriage rates across groups, only the changes in hours worked can generate the observed increase in the marriage rates experienced by high-income women.

This paper is broadly connected to the economic literature on marriage, which has gained significant traction since Becker's seminal works (Becker, 1973 and Becker, 1974). More specifically, it is closely related to research focusing on marriage trends and their respective driving factors. A particularly relevant study in this field is Greenwood et al. (2016), which develops a model to explain how technological advancements in home production and shifts in wage structure explain the drop in marriage rates from 1960 to 2005. As technology reduces household labor needs, more married women enter the workforce. This, accompanied by rising education returns and a narrowing gender wage gap, promotes higher education, results in an increase of positive assortative mating, and reduces the incentive for marriage. However, our analysis argues that while this mechanism successfully explains most marriage patterns from 1960 to 2005, it falls short of explaining why high-income women are marrying more today. Technological progress in home production, rather than increasing the incentive to marry, may instead enhance the appeal of remaining single.

Several studies have also identified shifts in the wage structure as a mechanism to explain the decline in marriage rates. Regalia and Ríos-Rull (2019), for instance, attributes a significant portion of this trend to changing relative wages both within and between sexes. Specifically, the study estimates that wage shifts account for roughly one-third of the rise in the share of single women and over half of the increase in single mothers from the mid-1970s to the early 2010s. This shift is largely driven by the reduction in the male-female wage gap, which diminished the financial benefits of marriage, especially for low-wage women.

Contrary to changes in wages, changes in work hours have not received serious attention when analyzing factors that shape marriage patterns.[1] Conversely, several recent papers suggest the potential impact that the decline in marriage rates might have on the reduction of hours worked by men. A recent hypothesis proposed by Binder and Bound (2019) postulates that declining marriage rates may contribute to the decrease in male work rates. Based on this, Blandin et al. (2023) suggests that recent declines in marriage rates, if exogenous, can account for a sizable share of the overall decline in prime-age male hours over recent decades. In contrast to their hypothesis, my analysis delves into how shifts in the quantity of work hours influence marriage formation in the last two decades, rather than how marriage formation affects work hours. To the best of my knowledge, this paper is the first to emphasize the critical role of work hours in such a context.

Additionally, other papers link structural change to declining marriage rates, suggesting that these shifts alter labor market stability and income expectations, which in turn influence marriage decisions. Anelli et al. (2021) find that in regions where the gender gaps in income and labor force participation are reduced due to exposure to robots, the relative economic status of men declines, as does the number of marriages. Autor et al. (2019) argue that rising international manufacturing competition has reduced employment and earnings of young adult males (in comparison to young adult females), consequently having a negative impact on marriages. These studies are among the few that emphasize the impact of men's economic outlook on marriage, which aligns with the narrative developed in this paper. I contribute to this line of research by demonstrating how a reduction in

[1] Among papers focusing on the intersection of research on marriage and time allocation are Knowles (2012) and Jones et al. (2015). Knowles (2012) examines intra-household bargaining and its effect on U.S. married men's labor supply from the mid-1970s to 2001. Using U.S. time-use data shows that bargaining slightly increased married men's weekly labor supply by 2.1 hours and decreased married women's by 2.7 hours. Jones et al. (2015) suggest that even slight decreases in gender wage disparities largely explain the increased work hours for married women, while technological gains in home production were less impactful.

men's working hours positively impacts the marriage rates of high-income women.

One notable paper by Newman (2018) documents the reversal in U.S. divorce rates, showing a negative correlation between female labor force participation and divorce rates. They argue that this trend is driven by high-income women forming more stable marriages, as dual-earner couples can compensate each other with *money and kind* (a balanced basket of market-procured and household-produced goods) an option less available to single earners or low-income couples. Building on this idea, I introduce the dimension of time, proposing that the time available to men, alongside financial stability, makes marriage more attractive to high-income women, thus reinforcing the trend of increased marriage rates within this group.

The remainder of the paper is structured as follows. The second section presents the empirical findings, the third sets the model, the fourth discusses the calibration, the fifth details the counterfactual exercises, and the sixth concludes the discussion.

4.2 Empirical findings

In this section, I present the empirical facts alongside a discussion of their relevance, specifically focusing on aspects overlooked in previous research. One missing piece in the literature on marriage trends relates to changes in the wage structure and work hours. Indeed, while many studies consider the reduction in the gender wage gap and its impact on marriage decisions, they overlook that this reduction was due not only to the sharp increase in female labor participation and subsequent gains in female wages but also to widening inequality among men.

I demonstrate this using IPUMS data from 2000 to 2021.[2] I divide my data into two income groups by gender. I define the aggregate median wage separately for men and women, using it as a threshold to categorize each gender into two distinct income groups: those earning below the median and those earning above it. For each income group, I calculate the median wage, providing a central measure of income for individuals within each category. I then compute the wage ratio for each income group, which is calculated as the group's respective median wage divided by the aggregate median wage of the gender. This approach allows for a meaningful comparison of wage dynamics within each gender over time, revealing variations in income growth or decline across different income groups of the male and female populations. Table 4.1 displays the ratio between the median wage of the income of the group and the aggregate median wage by gender for 2000 and 2021. Note that the only group performing slightly well for men is the second, with a 0.07% increase and a negative fall of 12% for group 1. For women, however, changes were positive with the highest increase of more than 5% corresponding to group 2.

In Table 4.2, I compute the female-to-male wage ratio using the median wage of each income group. I observe contrasting reductions in the wage gap, which suggests that

[2] I use the 1% American Community Survey and 5% samples from 2000 to 2021.

Table 4.1: Wage Ratio for Men and Women by Income Group

	2000	2021	Change (%)
Men			
Group 1 (poor)	0.647	0.567	**-12.36%**
Group 2 (rich)	1.416	1.417	**+0.07%**
Women			
Group 1 (poor)	0.546	0.548	**+0.36%**
Group 2 (rich)	1.512	1.595	**+5.89%**

Note: The wage ratio is computed as the median wage of each income group divided by the total median wage of each gender. Percent changes are calculated between the years 2000 and 2021.

women have fared well in the labor market between 2000 and 2021.

Table 4.2: Female-to-Male Wage Ratio for Income Quartiles

	2000	2021	Change (%)
α_1	0.558	0.676	**+21.15%**
α_2	0.706	0.788	**+11.61%**

Note: α_1 and α_2 represent the wage ratio between poor women to poor men, and rich women to rich men, respectively. Percent changes are calculated between the years 2000 and 2021.

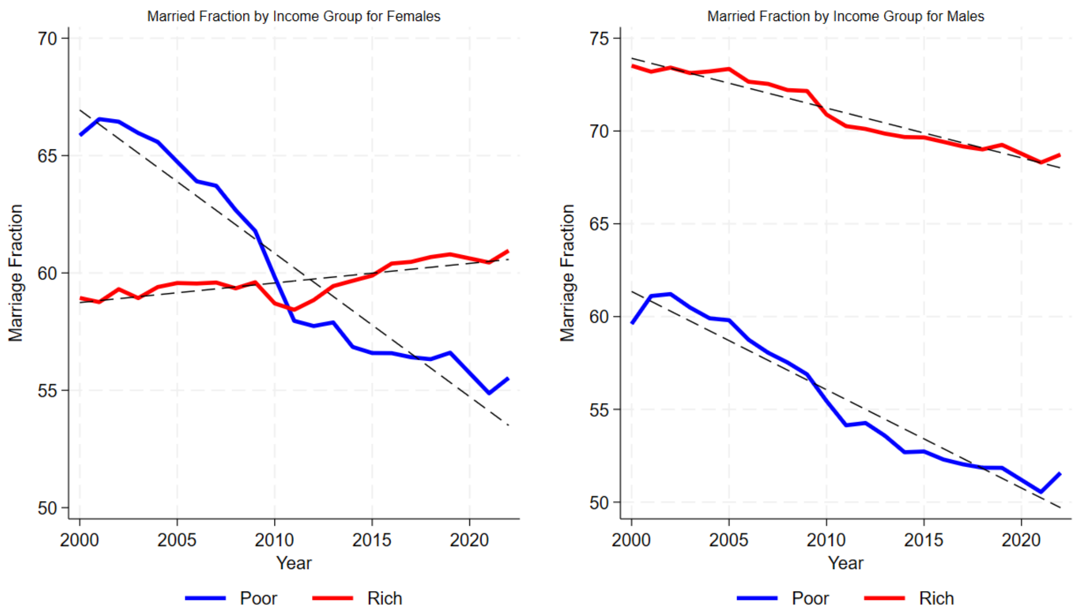
Using IPUMS data, I plot the married fractions (Figure 4.1) for two income groups for each gender.[3] It should be noted that the married fraction in Figure 4.1 and Figure 4.2 are for the population between 25 and 55 years of age.[4]

As can be seen in Figure 4.1, the results are striking. Initially, richer men married more in 2000 while poorer men married less. By 2021, the ranking remained similar although the marriage rates decreased for both groups, with a higher decline for the poorest. Conversely, for women, the opposite trend was observed; poorer women married more in 2000 while richer women married less. By 2021, this pattern reversed, with poorer women experiencing a decrease in marriage rates contrary to richer women, who had a marriage rate in 2021 higher than that in the first period. I add another graph with the

[3] I add as well, singlehood and cohabitation rates for four groups representing quartiles of income from the year 1960 to 2021 (Figures 4.6 and 4.7 respectively)See Appendix

[4] In the Appendix, for the graphs with four groups of income for each gender and longer period 1960-2021, I follow Chiappori et al. (2020) in selecting a sub-sample of individuals observed at or past 35 years of age and capping the age range at 44. This procedure minimizes the bias stemming from the fact that marriage takes place over a long age period and is most likely in later years of life, and I consider this as a robustness check.

Figure 4.1: Married Fraction for two Income Groups by gender



Note: Poor and rich are defined based on the median income for each gender. The married fraction represents the ratio of married individuals to the total population aged between 25 and 55 that is not married. Among all marital categories, only those who are married with a spouse present are considered married. The remaining categories (including married with spouse absent, separated, divorced, widowed, and never married/single) are classified as single.

same analysis with four quartiles of income for each gender. The output in Figure 4.2 is even clearer. The married fraction of the top quartile of women has been increasing from 2000 to 2021 while the married fraction of all remaining groups has been declining. For men, the richer the group the lower the decline. This observation represents a novel contribution to the existing body of literature.

I also analyze the patterns of sorting with the data since it includes unique identifiers for individuals and their spouses, which allows me to link married couples. Table 4.3 contains the percentage of each combination possible in the economy based on respective gender income groups.[5]

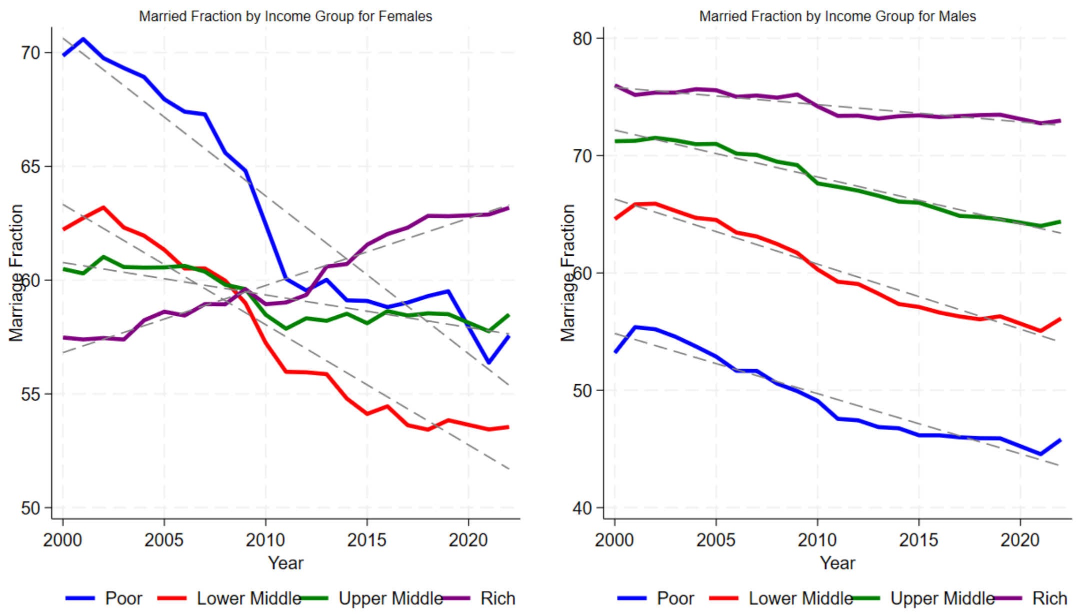
Table 4.3: Fraction of Couples

Year	(poor F,poor M)	(rich F, poor M)	(poor F,rich M)	(rich F,rich M)
2000	0.2612	0.1826	0.2528	0.3034
2021	0.2129	0.1733	0.2363	0.3775

Note: Each group (i,j) represents the pairing of individuals from income groups i and j in a given year. Poor and rich are defined based on the median income for each gender. Fractions of couples for each year sum to 1.

[5] The sum of all numbers in each year equals one.

Figure 4.2: Married fraction for four income groups by gender



Note: The income groups (poor, lower middle class, upper middle class, and rich) are defined based on income quartiles for each gender. The married fraction represents the ratio of married individuals to the total population aged between 25 and 55 that is not married. Among all marital categories, only those who are married with a spouse present are considered married. The remaining categories (including married with spouse absent, separated, divorced, widowed, and never married/single) are classified as single.

What is clear from Table 4.3 is that in the 2000s, marriages in which both husband is rich and the wife is poor accounted for 25 percent, whereas marriages with both rich husbands and wives represented 30 percent of the couples formed. In 2021, the first rate decreased by almost 5 percentage points and the second rate increased by almost 7 percentage points. To demonstrate that this trend started well before the 2000s, I include in the Appendix a contingency table of couple formation for the years 1960 and 2021 by income quartiles.[6]

The literature primarily focuses on educational levels to define assortative mating and study changes in marriage, but there may be more nuanced factors at play. Notably, while educational attainment is closely related to income levels, relying solely on education can

[6]Tables 4.10 and 4.11 show that in 1960, marriages in which the husband earns a higher income and the wife is a stay-at-home wife were the most prevalent, representing 19.8 percent of all couples that year. It should also be noted that groups 3 and 2 of men, who were predominantly married to non-working wives, represented 15.2 and 12.7 percent respectively. However, in 2021, the sorting patterns changed drastically with a more prevalent positive assortative mating in income. The highest percentage of couples involves rich men marrying rich women, comprising 11.6 percent of the total marriages, which is almost five times the rate in 1960. There is also an increase in the percentage of marriages on the diagonal. However, this table only indicates and does not prove assortative mating

be misleading. Over the last few decades, women have surpassed men in educational attainment, which has led to a larger share of women marrying men with lower education levels but higher incomes. Qian (2017) finds that educational assortative mating has reversed from a tendency for women to marry up in terms of education to a tendency for women to marry down, while the trend for women to marry men with higher incomes has persisted.

This is evident in the data from 1960, 2000, and 2021, showing that women indeed marry men with similar or higher incomes (see Tables 4.3 and 4.11). [7]

Table 4.4: Average Hours Worked Per Week by Income Quartile and Gender, 2000 and 2021

	2000	2021	Change (hours)
Men			
Group 1 (poor)	43.02	40.80	**-2.22**
Group 2 (rich)	46.92	45.25	**-1.67**
Women			
Group 1 (poor)	33.02	33.20	**+0.18**
Group 2 (rich)	41.55	41.99	**+0.44**

Note: This variable is defined in the census as the number of hours per week that the respondent usually worked if the person worked during the previous year.

In Table 4.4, I list the average hours worked per week for each gender and income group. A surprising result is the reduction in work hours for men from the year 2000 to 2021.[8] This reduction is substantial, given the short time period. Figure 4.3 illustrates how this variable changes for both men and women, with an increase in the average hours worked for both poor and rich men, and a slight increase for women from the year 2000 to 2021. This variation could impact marital decisions, as time is a crucial component of household utility.

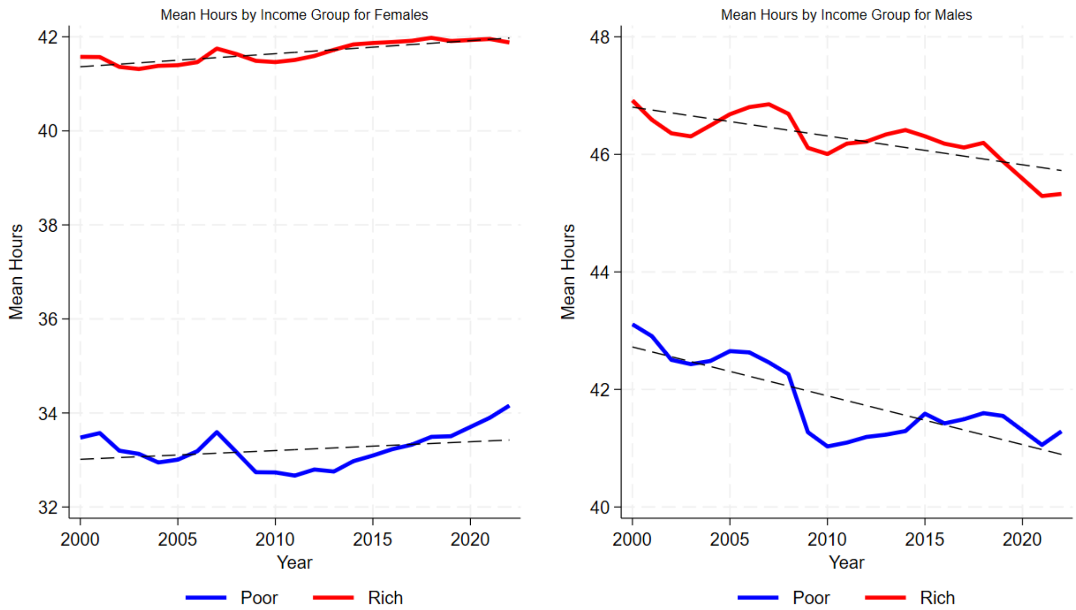
Consider this scenario: a man and a woman, both active in the labor market, decide whether to marry. If one partner reduces their work hours, it could significantly increase

[7] Based on these results, it is not unreasonable to suggest that the wage gap between these groups is a crucial determinant of marriage. This lends more credibility to the structural change narrative. Between 1950 and 1960, manufacturing served as the driving engine of growth and a major employer for men, especially those in the middle-income bracket. In contrast, women's participation in the labor force was relatively modest. With the decline of manufacturing and the rise of the service sector, a polarization of the labor force began (Autor and Dorn, 2013; Cerina et al., 2019), which disadvantaged men in manufacturing industries relative to women in the service sector.

[8] To put these numbers in perspective, the yearly reduction is more pronounced. Since the year comprises almost 52 weeks, I can observe that poorer and richer men are working 115.44 and 86.84 fewer hours, respectively, in 2021 than in 2000.

their likelihood of marrying, as the partner working less can contribute more to household duties or childcare, thereby offering more time to the family. If it is the man who reduces his work hours, then the likelihood of the woman wanting to marry him increases, and similarly, if the woman works less, the likelihood of the man wanting to marry her increases, assuming that wage levels remain constant.[9]

Figure 4.3: Mean Hours for two income groups by gender



Note: Poor and rich are defined based on the median income for each gender. The mean hours variable represents the number of hours per week usually worked in the previous year for each group. The variable covers the total population aged between 25 and 55 in each group.

In this context, and reflecting the observations in the data, a reduction in men's work hours could lead to an increased probability of women desiring to marry. This explanation is supported by Doepke et al. (2022) which asserts that traditional notions about fertility are evolving, with the negative relationship between women's labor force participation and fertility becoming positive in the developed world. They identify cooperative fatherhood as a likely factor in this shift. This aligns closely with our argument that men working fewer hours could be more appealing to women in the marriage market, assuming that a reduction in work does not equate to lower earnings.

As the main purpose of this paper is to determine which of these two channels, wage and time, better explains the recent marriage trends observed in the US over the last 22 years, in the next section I construct a marriage model that incorporates wage and time as the principal components influencing the decision to marry.

[9] I add the same graph for four quartiles in the Appendix (Figure 4.4)

4.3 Model

I focus on two economic factors: wage disparities within and between genders, and changes in working hours. For wages, the notable shifts are the reduction of the wage gap between men and women and the widening wage gap between rich and poor men. These shifts lead to two significant outcomes. First, the reduction in the gender wage gap makes women less likely to pursue marriage, as the value of remaining single increases. Second, the increasing wage gap between rich and poor men negatively impacts the marriage prospects of poorer men. Consequently, I observe fewer marriages and more assortative matching, where richer men are more likely to marry richer women. Additionally, a reduction in working hours for men makes them more attractive to women, as it allows for more time for activities both within and outside the household.

Two key components are needed to formalize the discussion above. First, a model of marriage is essential. Second, there must be heterogeneity in wages and working hours among individuals. This forms the basis for the following setup.

Let us assume an economy with a unitary mass of male and female agents. I also assume that men and women belong to two income groups (rich and poor). In this model, a person can only get married or remain single as I assume "happy-ever-after" marriages without divorce. Every individual faces a constant probability of dying, δ, each period. Following Greenwood et al. (2016), any dead individual is replaced by a single doppelganger of the same type. I include the probability of dying, in each period to ensure a stationary distribution of agents over time. This simplifies the analysis by allowing us to focus on steady-state equilibrium without worrying about population growth or decline.

Agents search for partners while being single, and a couple is only formed if both parties mutually agree to marry. I assume that couples derive utility solely from pooled income and non-working time. The significance of non-working time in marriage decisions is supported by the findings of Aguiar et al. (2013), which demonstrates that adjustments in time use in response to changes in market work differ significantly between married individuals and singles. Specifically, within the married group, approximately 42 percent of the hours not spent on market work are redirected to non-market activities and child care. This indicates that married individuals place substantial value on non-market time, utilizing it for home production and family-related activities. Although our model does not explicitly include fertility decisions or child-rearing activities, incorporating non-working hours into the utility function effectively captures their impact on marriage decisions. By valuing non-market time, I acknowledge that individuals derive utility not only from consumption but also from activities outside of market work, such as leisure, home production, and family care. This inclusion allows the model to reflect observed behaviors where reductions in working hours can increase the attractiveness of marriage for women (or otherwise for men), by enabling partners to allocate more time to shared non-market activities. Therefore, the utility derived from non-working time plays a crucial role in influencing individuals' decisions to marry, as it encompasses the benefits of spending time together, raising children, and engaging in household production.

I base the model on some elements of Greenwood et al. (2016). Mainly, I assume that the total non-working time available for non-market activities is given by:

$$n = \begin{cases} 1 - h^s_{g,i}, & \text{for } \textit{singles}, \\ 2 - h^m_{g,i} - h^m_{g*,i*}, & \text{for } \textit{married}, \end{cases} \tag{4.1}$$

where $h^k_{g,i}$ and $h^k_{g*,i*}$ are the total amount of time spent on the labor market for gender $g(g*)$ with income level $i(i*)$ and marital status $k = \{s, m\}$. It is important to note that labor market participation is treated as an exogenous variable in this model. Each group is assigned the corresponding hours based on marital status. The variable n represents everything that can be produced by non-worked time, including home production, leisure, child care, and other non-market activities.

I depart from Greenwood et al. (2016) by abstracting from the home goods production function and focusing solely on non-working hours in the market. This simplification is intentional, as I argue that the impact of household durables prices on marriage dynamics has diminished since the 2000s.[10]

The following equations describe the instantaneous utility of both singles, u_s, and married individuals, u_m:

$$u_s(c, n) = \frac{1}{1-\zeta}(c - \mathfrak{c})^{1-\zeta} + \frac{\alpha}{1-\xi} n^{1-\xi}, \tag{4.2}$$

$$u_m(c, n) = \frac{1}{1-\zeta}\left(\frac{c - \mathfrak{c}}{\chi}\right)^{1-\zeta} + \frac{\alpha}{1-\xi}\left(\frac{n}{\chi}\right)^{1-\xi}, \tag{4.3}$$

with

$$c = \begin{cases} w^s_{g,i}, & \text{for } \textit{singles}, \\ w^m_{g,i} + w^m_{g*,i*}, & \text{for } \textit{married}, \end{cases} \tag{4.4}$$

[10] Greenwood et al. (2016) found that the price of household durables relative to other goods decreased by about 60 percent between 1960 and 2005, significantly influencing marriage decisions during that period. However, this trend has not continued at the same pace in recent decades. Specifically, in the National Income and Product Accounts (NIPA), the price index for "furnishings and durable household equipment" relative to the price index for "personal consumption expenditures" dropped by about 35 percent between 2000 and 2020, which is still significant but a smaller decline compared to earlier periods. Moreover, as highlighted by Church (2014), over the last three decades, consumer expenditures have shifted from commodities, including durable goods, to services, particularly housing. The analysis shows a significant increase in the quantity of housing due to factors such as higher homeownership rates and larger home sizes, while the quantity of durable goods has increased modestly. This shift indicates that the role of household durables in influencing marriage decisions might have become less significant in the modern context. I do not explore the role of services or the potential impact of rising service prices on marriage dynamics within this model. However, the increase in service prices could help explain positive assortative matching in income, as couples may require a high combined income to afford these services, especially housing.

where $\mathfrak{c}$ is the minimum consumption requirement and χ is the household equivalence scale. The parameter $\zeta > 0$ determines the curvature of the utility function with respect to market goods consumption (c). Similarly, the parameter $\xi > 0$ determines the curvature of the utility function with respect to non-market time (n).

Next, I denote the value of being single as $V_{g,s}(w^s_{gi})$, and the value of being married with the assumption of no divorce as $V_{g,m}(w_{gi}, w_{g^*i^*})$. Note that the values $V_{g,s}$ and $V_{g,m}$ also depend on hours worked. Since these hours are group-specific and linked to specific wages, the functions are written here as depending primarily on wages for simplicity. Each agent draws a match-specific shock ϵ_k drawn from cumulative density function (CDF) F^ϵ which has a continuous domain and is a Type I extreme value distribution. If a couple $(w_{g,i}, w_{g^*,i^*})$ meets, and given the extreme value shocks, the probability of that agent g wants to marry is:[11]

$$D_g(w_{gi}, w_{g^*i^*}) = \frac{\exp\left(\frac{V_{g,m}(w_{gi},w_{g^*i^*})}{\sigma_\epsilon}\right)}{\exp\left(\frac{V_{g,m}(w_{gi},w_{g^*i^*})}{\sigma_\epsilon}\right) + \exp\left(\frac{V_{g,s}(w_{gi})}{\sigma_\epsilon}\right)} \tag{4.5}$$

where σ_ϵ is the scale parameter of the extreme value shock. This can be seen as the love parameter.[12] The expression in 4.5 represents the probability of the utility from marrying exceeds the utility from remaining single.

A couple is formed if both individuals want to do it. That is, two individuals will get married only if their corresponding utilities when married are higher than when single. Hence, the probability of a couple marrying is given by the product of both probabilities D_g and D^*_g. I denote this product by:

$$\Lambda_{gg^*}(.) = \text{Probability of marriage} = D_g(.) \times D_{g^*}(.) \tag{4.6}$$

Using the matching rule Λ_{gg^*}, I derive the value of being single as:

$$\begin{aligned} V_{g,s}(w^s_{gi}) = u_s(w^s_{gi}) + \beta \Bigg\{ p \left[\sum_{w^s_{g^*i^*} \in W} \frac{M_{g^*,s}(w_{g^*i^*})}{M_{g^*,s}} \tilde{V}_{g,m}(w_{gi}, w_{g^*i^*}) \right] \\ + (1-p) V_{g,s}(w_{gi}) \Bigg\} \end{aligned} \tag{4.7}$$

with

$$\begin{aligned} \tilde{V}_{g,m}(w^m_{gi}, w^m_{g^*i^*}, w^s_{gi}) = [\Lambda_{gg^*}(w_{gi}, w_{g^*i^*})] V_{g,m}(w^m_{gi}, w^m_{g^*i^*}) \\ + [1 - \Lambda_{gg^*}(w_{gi}, w_{g^*i^*})] V_{g,s}(w^s_{gi}) \end{aligned} \tag{4.8}$$

with β representing the discount factor, p representing the probability of a single individual meeting a potential partner in a given period, $M_{g^*,s}(w_{g^*,i^*})$ representing the measure of single agents of the opposite gender g with wage w_{g^*,i^*} and $M_{g^*,s}$ representing the total

[11] The derivation of D_g can be found in Appendix A.1.

[12] Note that when σ goes to 0, D goes to 1 if $V_{g,m} > V_{g,s}$ and zero otherwise, translating purely the models power. On the other hand, when σ goes to 1, the probability goes to 0.5.

measure of single agents of the other agents g^*. Thus, the ratio between $M_{g^*,s}(w_{g^*,i^*})$ and $M_{g^*,s}$ is the probability that a potential partner from the opposite gender has a wage (w_{g^*,i^*}).

Equation 4.8 is the expected value of transitioning from being single to being married, accounting for the probability of marrying Λ_{gg^*}. The intuition behind 4.7 mirrors the timing of events in the model. A single individual goes to the marriage market, he remains single with a probability of $(1-p)$, as such the expected value if the individual does not meet a potential partner is $(1-p)V_{g,s}(w_{gi})$. There is a probability p that he finds a potential partner. Moreover, since partners of the opposite gender could either be rich or poor, the probability of finding anyone of them will depend on the measures of these groups. For the sake of simplicity, let us assume that the individual is a poor woman, she finds a poor man. This does not necessarily mean they will end up marrying. Depending on the matching rule, she might marry, with probability $\Lambda_{fm}(w_{fp}, w_{mp})$, or choose to remain single with probability, $1-\Lambda_{fm}(w_{fp}, w_{mp})$.

As for the value of being married, I have:

$$V_{g,m}(w_{gi}^m, w_{g^*i^*}^m) = u_m\left(w_{gi}, w_{g^*i^*}\right) + \mu_{i*}^g + \beta V_{g,m}(w_{gi}, w_{g^*i^*}) \tag{4.9}$$

where μ_{i*}^g is a parameter that represents the preference of agents of the same gender towards the income group i of the opposite gender. It can be interpreted as a preference parameter capturing the additional utility or disutility that an individual of gender g derives from being married to a partner of income group i of the other gender. This might reflect social preferences and can be interpreted as status consideration.

Solving the single's problem requires knowing the steady-state single and married population. The measure of singles is defined by:

$$\begin{aligned} M_{g,s}(w_{gi}) = (1-\delta)M_{g,s}(w_{gi}) \left\{ p \left[\sum_{w_{g^*i^*} \in W} \frac{M_{g^*,s}(w_{g^*i^*})}{M_{g^*,s}} [1-\Lambda_{gg^*}(w_{gi}, w_{g^*i^*})] \right] + (1-p) \right\} \\ + \delta(M_{g,s}(w_{gi}) + M_{g,m}(w_{gi}, w_{g^*i^*})) \end{aligned} \tag{4.10}$$

As equation (4.10) describes, the number of singles today depends on surviving singles from the previous period who did not marry, doppelgangers replacing dead married and single individuals.

Similarly, the measure of married agents is defined by:

$$\begin{aligned} M_{g,m}(w_{gi}, w_{g^*i^*}) = (1-\delta)M_{g,s}(w_{gi})p\frac{M_{g^*,s}(w_{g^*i^*})}{M_{g^*,s}}\Lambda_{gg^*}(w_{gi}, w_{g^*i^*}) \\ + (1-\delta)M_{g,m}(w_{gi}, w_{g^*i^*}) \end{aligned} \tag{4.11}$$

equation (4.11) describes the dynamics of the married population. The number of married individuals today depends on the surviving married individuals and surviving singles who get married.

A stationary matching equilibrium is a set comprising the value function for single agents $V_{g,s}(w^s_{gi})$, the value function for married agents $V_{g,m}(w^m_{gi}, w^m_{g^*i^*})$, a decision rule $D_g(w_{gi}, w_{g^*i^*})$, a matching rule $\Lambda_g(w_{gi}, w_{g^*i^*})$, and stationary distributions for the measures of singles and married couples $M_{g,s}(w_{gi})$ and $M_{g,m}(w_{gi}, w_{g^*i^*})$, for each gender $g = \{f, m\}$ and income group $i = \{poor, rich\}$.

(I) The value function $V_{g,s}(w^s_{gi})$ solves the single's recursion (Equation 4.7), taking as given the instantaneous utility function for singles $u_s(w^s_{gi})$ from Equation 4.2, the value function for married individuals $V_{g,m}(w^m_{gi}, w^m_{g^*i^*})$ from 4.9, the matching rule $\Lambda_g(w_{gi}, w_{g^*i^*})$ 4.6, and the stationary distribution for the measure of singles $M_{g,s}(w_{gi})$ in 4.10.

(II) The value function $V_{g,m}(w^m_{gi}, w^m_{g^*i^*})$ solves Equation (4.9), taking as given the indirect instantaneous utility function for married individuals $u_m(w^m_{gi}, w^m_{g^*i^*})$ from Equation 4.3.

(III) The decision rule $D_g(w_{gi}, w_{g^*i^*})$ solves the marriage decision problem (Equation 4.5), taking as given the value functions $V_{g,s}(w^s_{gi})$ and $V_{g,m}(w^m_{gi}, w^m_{g^*i^*})$.

(IV) The matching rule $\Lambda_g(w_{gi}, w_{g^*i^*})$ is solved based on Equation 4.6, taking as given the decision rules $D_g(\cdot)$ and $D_{g^*}(\cdot)$.

(V) The stationary distributions $M_{g,s}(w_{gi})$ and $M_{g,m}(w_{gi}, w_{g^*i^*})$ solve Equations 4.10 and 4.11, taking as given the matching rule $\Lambda_g(\cdot)$.

4.4 Calibration

I calibrate the parameters of the model to the 2000 US data. Some parameters are chosen and assigned based on a priori information. In particular, I assume that a period is one year and I set delta $\delta = 0.033$, which matches the death probability of individuals aged 25 to 55. Consequently, the discount factor satisfies $\beta = 0.960 \times (1 - 0.033)$. I assign the value for the work week by taking the average hours worked for each group from the data and dividing it by 112, which represents non-sleeping hours in a week. The equivalence scale is set at $X = 0.70$ from the (OECD). Finally,$\{\zeta, \bar{c}\}$ are taken from Greenwood et al. (2016). Table 4.5 summarizes these parameters.

The parameters to calibrate are $\{\mu^m_p, \mu^m_r, \mu^f_p, \mu^f_r, \sigma, \alpha, \xi\}$. These parameters are estimated such that the model matches 8 data moments for the year 2000. The data targets are the fraction of the population that has been married by gender and income group,

$$\frac{M_{g,m}(w, w^*)}{M_{g,m}}$$

and the fraction of couples for each combination of the income types for both the husband and wife.

$$\frac{\sum M_{g,m}(w, w^*)}{\sum M_{g,s}(w^*) + M_{g,m}(w, w^*)}$$

Table 4.5: Parameters: A Priori Information

Category	Parameter values	Criteria
Preferences	$X = 0.70$	OECD scale
	$\beta = 0.96$	Prescott (1986)
	$\zeta = 1.782$	Greenwood et al. (2016)
	$\bar{c} = 0.068$	Greenwood et al. (2016)
Death probability	$\delta = 1/30$	A 30-year lifespan
Hours	$h_{m1,m,2000} = 0.3842$	Data
	$h_{m2,m,2000} = 0.4209$	Data
	$h_{f1,m,2000} = 0.2969$	Data
	$h_{f2,m,2000} = 0.3713$	Data
	$h_{m1,s,2000} = 0.3669$	Data
	$h_{m2,s,2000} = 0.4131$	Data
	$h_{f1,s,2000} = 0.3286$	Data
	$h_{f2,s,2000} = 0.3810$	Data
Wages	$w_{f2,s,2000} = 1$	(normalization)
	$w_{m1,m,2000} = 1.000$	Data
	$w_{m2,m,2000} = 2.114$	Data
	$w_{f1,m,2000} = 0.512$	Data
	$w_{f2,m,2000} = 1.463$	Data
	$w_{m1,s,2000} = 0.858$	Data
	$w_{m2,s,2000} = 2.033$	Data
	$w_{f1,s,2000} = 0.569$	Data
	$w_{f2,s,2000} = 1.455$	Data

Note: The table displays the median wages for each income group, normalized by the median wage of high-income women. This wage data, sourced from the IPUMS ACS, represents pre-tax wages and salary income received as an employee during the previous year. The hours listed indicate the number of hours worked per week and are drawn from the same source as the wage data. In the subscripts for wages and hours, the first letter denotes the gender, and the second letter represents their marital status (married or single).

Table 4.6 reports the parameter estimates. The set of moments and the corresponding results for the benchmark model for the year 2000 are displayed in Table 4.7. The fitted parameter values match the data very closely.

Table 4.6: Parameters: Estimated

Category	Parameter	Value
Preferences	μ_{m1}	-0.354754119334353
	μ_{m2}	-0.212772033060683
	μ_{f1}	-0.453297033494276
	μ_{f2}	-0.526434406793840
	σ	1.300231463477392
	α	2.100429049320298
	ξ	2.065317083422249

Nonmarket time is calibrated to have a weight of $\alpha = 2$ in the utility. This is equivalent to a weight of leisure amounting to 0.9, in a typical macro model with a consumption weight equal to 0.45 ($0.9/0.45 = 2$). In Greenwood et al. (2016) α is 1.20, which in this context translates to a weight of 0.55 being applied to leisure. Note that our estimate is higher and this is consistent with the idea that couples care more about leisure in the 2000s than they did in 1960, and thus it is not unreasonable. The utility function for "nonmarket time" is less concave than Greenwood et al. (2016) ($\xi_{our} = 2.06 < \xi_{2016} = 3.11$) but still greater than the utility's degree of curvature of market goods ($\zeta = 1.78$). This implies that more household wealth leads to more allocation toward market goods, though with a lesser degree than it is in the estimate ofGreenwood et al. (2016). The type preferences μ capture gender's heterogeneity with respect to marriage.

Going back to Table 4.7, the model matches most of the targets with ease. It is worth noting, however, that the model is less effective in matching two moments: The fraction of rich husbands married to poor wives and the fraction of poor husbands married to rich wives. The data sets the first moment at 25.62 percent whereas the model predicts it to be 27.49 percent. For the second moment, the data has it at 18.26 percent, whereas the model underestimates it at 16.05 percent.

4.5 Moving to 2021

For the purpose of simulating the model for the year 2021, I use the 2021 wages and hours worked.[13] In this section I assess the model's fit to the US data for the year 2021.

[13] 2021 Wages are deflated to be comparable to the 2000 dollar values. Wages are divided by the 2021 consumer price index and then multiplied by the 2000 consumer price index. I normalize by the wage of poor men in 2000. After this process, we note a decrease in the wage of poor married and single men and a growth for the rest of the groups. I also notice a widening gap

Table 4.7: Data and Benchmark Model, year 2000

Category	Income groups	Data	Model
Marriage Rates	Females (poor)	0.6582	0.6564
	Females (rich)	0.5894	0.5937
	Males (poor)	0.5951	0.5976
	Males (rich)	0.7348	0.7307
Sorting	Poor Wife - Poor Husband	0.2612	0.2562
	Poor Wife - Rich Husband	0.2562	0.2749
	Rich Wife - Poor Husband	0.1826	0.1605
	Rich Wife - Rich Husband	0.3033	0.3083

Note: Poor and rich are defined based on the median income for each gender. The married rates represent the ratio of married individuals to the total population aged between 25 and 55 that is not married. Among all marital categories, only those who are married with a spouse present are considered married. The remaining categories (including married with spouse absent, separated, divorced, widowed, and never married/single) are classified as single. Sorting indicates the income pairing between wives and husbands of the different income groups.

In the next section, I execute a decomposition exercise to assess the importance of the two mechanisms at hand: the change in the wage structure and the change in the hours worked. Running the model with the wage structure and hourly work of the year 2021, I get the result in Table 4.8.

The main result is that the model successfully reproduces the trends observed in the data, specifically a decrease in marriage rates for all groups except for rich women. This is a new and relevant finding. To the best of my knowledge, this is the first paper to provide a model that manages to explain the contrasting differences in marriage trends across different groups.

With regards to marriage rates, the model manages to capture a sizable fraction of the change from the year 2000 to 2021. The model explains 47 percent of the decrease in the marriage rate of poor men, 74 of the decrease in the marriage rates of poor women, and 34 percent of the decrease in the marriage rates of rich men. As for rich women, the model overestimates slightly the marriage rates explaining 131 percent of the change.

Looking at the sorting contingency tables, two noticeable conclusions arise. First, the model does poorly when it comes to predicting the fraction of rich husbands to poor wives. This moment decreases in the data from 0.2562 in 2000 to 0.2410 in 2021, whereas it increases in the model from 0.2749 in 2000 to 0.2931 in 2021. The second is the fact that the model overestimates to some extent the fraction of poor husbands to poor wives. This moment decreases in the data from 0.2612 in 2000 to 0.2131 in 2021, whereas it decreases

between married and single rich women in 2021 compared to 2000.

in the model from 0.2562 in 2000 to 0.1569 in 2021. This is due to the strength of the wage mechanism. Poor men had a decline in their wage from 2000 to 2021 in comparison, which reduced their value function and the value function of women who marry with their type.

Table 4.8: Data and Benchmark Model, Year 2000 and 2021

Category	**Income groups**	**Year**	**Data**	**Model**
Marriage Rates	Females (poor)	2000	0.6582	0.6562
	Females (rich)	2000	0.5894	0.5937
	Males (poor)	2000	0.5951	0.5976
	Males (rich)	2000	0.7348	0.7307
	Females (poor)	2021	0.5487	0.6070
	Females (rich)	2021	0.6045	0.6093
	Males (poor)	2021	0.5055	0.5291
	Males (rich)	2021	0.6829	0.7172
Sorting	Poor Wife - Poor Husband	2000	0.2612	0.2562
	Poor Wife - Rich Husband	2000	0.2562	0.2749
	Rich Wife - Poor Husband	2000	0.1826	0.1605
	Rich Wife - Rich Husband	2000	0.3033	0.3083
	Poor Wife - Poor Husband	2021	0.2131	0.1569
	Poor Wife - Rich Husband	2021	0.1808	0.1841
	Rich Wife - Poor Husband	2021	0.2410	0.2931
	Rich Wife - Rich Husband	2021	0.3649	0.3657

Notes: Data and Model columns represent empirical data and the corresponding benchmark model values for the years 2000 and 2021. Poor and rich are defined based on the median income for each gender. The married fraction represents the ratio of married individuals to the total population aged between 25 and 55 that is not married. Among all marital categories, only those who are married with a spouse present are considered married. The remaining categories (including married with spouse absent, separated, divorced, widowed, and never married/single) are classified as single. Sorting indicates the income pairing between wives and husbands of the different income groups.

4.6 Decomposition

In this section, I will inspect our two main mechanisms. The change in the wage structure and the change in the hours worked. In this context, I consider two experiments. First, I attempt to see the impact of the change in the wage structure while keeping hours worked at the 2000 level. Second, I turn off the shifts in the wage structure and apply the changes in the hours worked. For this purpose, I use the results in Table 4.8 as the benchmark for the experiments. First, I consider the situation in which wages by gender and marital status are fixed to the 2000 levels and change only hours worked.[14] I see the results of this exercise in Table 4.9 under the column of experiment 1. I can clearly distinguish that the marital rate for rich women increases compared to the 2000 levels. On the other hand, I get a modest decrease in the rate of poor women and a weak decrease in the marital rates of poor men. Rich men also had a very small decrease.

Second, I consider the situation in which the hours by gender and marital status are fixed to the 2000 levels and apply the wage structure corresponding to the year 2021. I see the results of this exercise in Table 4.9 under the column of experiment 2. Interestingly, I get a decrease in the marital rate of rich women (0.5937 in the benchmark versus 0.5875 when only wage is changing). This outcome is expected, as the reduction in the wage gap between high-income men and women diminishes the financial benefits of marriage for women, according to this mechanism. One noteworthy point is that the impact of the closing gap between rich men and women is not offset by the widening gap between rich and poor women. As for the other marital rates, we see a sharper decrease for poor men and women and a weak decrease for rich men.

Table 4.9: Data and Benchmark Model, 2000 and 2021

	Benchmark		Experiments (2021)	
	2000	2021	Exp 1	Exp 2
Females (poor)	0.6562	0.6070	0.6279	0.6329
Females (rich)	0.5937	0.6093	0.6049	0.5794
Males (poor)	0.5937	0.5291	0.5767	0.5593
Males (rich)	0.7307	0.7172	0.7296	0.7262

Notes: Benchmark columns represent model values for the years 2000 and 2021. The numbers are the fractions of married. Exp 1 is the counterfactual with hours changes only, and Exp 2 is the counterfactual with wage changes only. Poor and rich are defined based on the median income for each gender. The married fraction represents the ratio of married individuals to the total population aged between 25 and 55 that is not married. Among all marital categories, only those who are married with a spouse present are considered married. The remaining categories (including married with spouse absent, separated, divorced, widowed, and never married/single) are classified as single.

These two experiments suggest that while the reduction in the wage gap negatively impacts marital rates across all groups, changes in hours worked do not. This indicates that time allocation makes marriage more appealing for higher-income women, possibly

[14]Note that I only change the wage structure and hours in this counterfactual exercise. I keep population counts for each group at their 2000 level. Although this paper is not concerned with explaining the increase in assortative matching, changing the population distribution across the different groups can indeed explain the phenomenon. This mirrors the increase of the stock of educated and working females in other papers of the literature

because husbands with more time can help with household production.

4.7 Conclusion

This paper investigates the recent trends in marriage rates in the United States from 2000 to 2021, focusing on the roles of wage structure and work hours. Utilizing IPUMS data, the study highlights that while changes in the wage structure can explain the decrease in marital fractions, a significant reduction in work hours for men across all income groups, contrasted with a slight increase for women, played a crucial role in influencing marriage trends.

The analysis reveals that reductions in male work hours may increase the probability of marriage, underscoring the importance of time availability in household dynamics. A calibrated marriage model tested with 2021 data supports the hypothesis that changes in work hours better explain recent marriage trends than changes in wages.

Through counterfactual exercises, the study demonstrates that the 2021 wage structure does not align with observed trends in marriage rates, particularly for rich women. In contrast, the model incorporating 2021 hours worked explains why high-income women marry more.

These findings help us better understand the economic factors that influence marriage, especially highlighting the importance of work hours alongside wages. They suggest that policies focused on improving work-life balance could significantly impact marriage trends, particularly as the job market continues to evolve. For example, introducing flexible work schedules or promoting remote work options could enable people to better juggle their professional and personal lives. These changes might reduce the stress and time constraints associated with long work hours, making marriage a more appealing and achievable option for those who might otherwise delay or forgo it due to work commitments. Additionally, enhancing parental leave policies could support couples in balancing family responsibilities with their careers, potentially leading to higher marriage rates and more stable family structures.

The model can be extended to incorporate a work-life balance policy. I expect that such policies would positively influence marriage rates by easing the economic and time pressures that often deter individuals from getting married and, hence, strengthen the mechanism in the model. By providing greater time flexibility, these policies could encourage earlier and more stable marital commitments, thereby shifting existing marriage trends. Moving forward, I plan to integrate specific work-life balance policies into my economic models to thoroughly assess their impact on marriage patterns. This includes evaluating how different policy measures perform and understanding their interactions with changes in the wage structure and other factors. Addressing these areas is a key part of my agenda, as I aim to develop a more comprehensive and nuanced understanding of the factors driving marriage trends in today's society.

Appendix C

A.1 Utility Maximization and Choice Probabilities

In this section, I explain how an agent's probability of marrying arises from utility maximization within the discrete choice framework. I derive the expression for D_g by modeling agents' decision-making processes under uncertainty due to unobserved preference shocks.

Each agent g faces a choice between two alternatives, marrying a potential partner denoted by m) or remaining single(denoted by s).

The utility associated with each choice has two components represented by the deterministic utility component which depends on observable characteristics such as (w_{gi}) and work hours (h_{gi}) in our model, as well as a random utility component which captures unobserved factors affecting the agent's preferences such as *love*.

The utility functions that the agent g seeks to maximize are specified as:

$$U_{g,m} = V_{g,m}(w_{gi}, w_{g^*i^*}, h_{gi}, h_{g^*i^*}) + \epsilon_{g,m}, \tag{4.12}$$

$$U_{g,s} = V_{g,s}(w_{gi}, h_{gi}) + \epsilon_{g,s}, \tag{4.13}$$

where $V_{g,m}$ is the deterministic utility of marrying. $V_{g,s}$ is the deterministic utility of remaining single. $\epsilon_{g,m}$ and $\epsilon_{g,s}$ are the random utility components (preference shocks).

I assume that the preference shocks $\epsilon_{g,m}$ and $\epsilon_{g,s}$ are independently and identically distributed (i.i.d.) according to a Type I Extreme Value distribution (Gumbel distribution):

$$\epsilon_{g,m} \sim \text{Type I Extreme Value}, \tag{4.14}$$

$$\epsilon_{g,s} \sim \text{Type I Extreme Value}. \tag{4.15}$$

This assumption facilitates the derivation of a closed-form expression for the choice probabilities.

Agent g will choose to marry if the utility from marrying exceeds the utility from remaining single $U_{g,m} > U_{g,s}$. The probability that agent g chooses to marry is then:

$$D_g = \Pr(U_{g,m} > U_{g,s}). \tag{4.16}$$

Substituting the utility functions:

$$\begin{aligned} D_g &= \Pr\left(V_{g,m} + \epsilon_{g,m} > V_{g,s} + \epsilon_{g,s}\right), \\ &= \Pr\left(\epsilon_{g,s} - \epsilon_{g,m} < V_{g,m} - V_{g,s}\right). \end{aligned} \tag{4.17}$$

Since $\epsilon_{g,m}$ and $\epsilon_{g,s}$ are i.i.d. Type I Extreme Value, the difference $\epsilon_{g,s} - \epsilon_{g,m}$ follows a logistic distribution. The cumulative distribution function (CDF) of the logistic distribution is:

$$F(z) = \frac{1}{1 + \exp\left(-\dfrac{z}{\sigma_\epsilon}\right)}, \tag{4.18}$$

where σ_ϵ is a scale parameter.

Therefore, the probability that agent g chooses to marry is:

$$\begin{aligned} D_g &= \Pr\left(\epsilon_{g,s} - \epsilon_{g,m} < V_{g,m} - V_{g,s}\right), \\ &= F\left(V_{g,m} - V_{g,s}\right), \\ &= \frac{1}{1 + \exp\left(-\dfrac{V_{g,m} - V_{g,s}}{\sigma_\epsilon}\right)}, \\ D_g &= \frac{\exp\left(\dfrac{V_{g,m}}{\sigma_\epsilon}\right)}{\exp\left(\dfrac{V_{g,m}}{\sigma_\epsilon}\right) + \exp\left(\dfrac{V_{g,s}}{\sigma_\epsilon}\right)}. \end{aligned} \tag{4.19}$$

The choice probabilities reflect the likelihood that the utility of marrying exceeds that of remaining single. An increase in the deterministic utility of marrying ($V_{g,m}$) relative to remaining single ($V_{g,s}$) increases the probability D_g that the agent chooses to marry. As for the scale parameter σ_ϵ influences the impact of the random utility components. A smaller σ_ϵ implies less variance in the unobserved preference shocks, making choices more deterministic. Whereas, a larger σ_ϵ increases the randomness in choices due to greater influence from unobserved factors.

A.2 Additional: Model

$$m(M_{ms}, M_{fs}) = \frac{M_{ms} \cdot M_{fs}}{\left(M_{ms}^p + M_{fs}^p\right)^{\left(\frac{1}{p}\right)}} \tag{4.20}$$

where M_{ms} is the mass of single males in the marriage market, and M_{fs} is the mass of single females in the marriage market. The matching function, $m(M_{ms}, M_{fs})$, represents the effective or "matched" mass when single males and females interact in the marriage market. The numerator, $M_{ms} * M_{fs}$, suggests that the matched mass increases with both the mass of single males and single females. The denominator introduces a non-linearity, controlled by the parameter p. This parameter determines the sensitivity of the matching function to changes in the masses of single males and females. If $p = 1$, the matching function simplifies to the geometric mean of M_{ms} and M_{fs}. p increases, the function becomes more sensitive to the smaller of the two masses, emphasizing the importance of balance in the market

$$p_m(M_{ms}, M_{fs}) = \frac{m(M_{ms}, M_{fs})}{M_{ms}} \tag{4.21}$$

This function calculates the proportion of the matched mass attributed to single males. It provides insight into the relative "influence" or "contribution" of single males to the overall matched mass in the marriage market.

$$p_f(M_{ms}, M_{fs}) = \frac{m(M_{ms}, M_{fs})}{M_{fs}} \tag{4.22}$$

Similarly, this function calculates the proportion of the matched mass attributed to single females. It offers a perspective on the relative ”influence” or ”contribution” of single females to the overall matched mass in the marriage market.

A.3 Data

Variables as defined in IPUMS-USA (Ruggles et al. (2024)):

- **MARST** indicates each person's marital status (Married, spouse present; Married, spouse absent; Separated; Divorced; Widowed; Never married/single).

- **UHRSWORK** reports the number of hours per week the respondent usually worked in the previous year, based on different reference periods in census surveys. It's a two-digit variable, with codes for missing, edited, or not applicable data, adjusted per census year and sample.

- **SEX**: Indicates the respondent's gender (male or female).

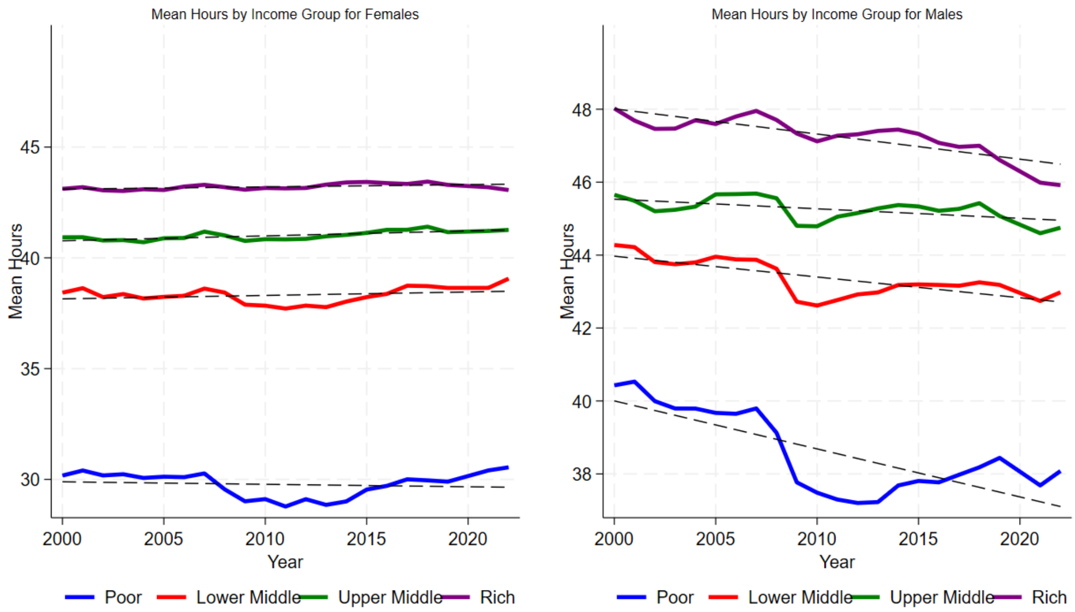
- **INCWAGE**: Reports total pre-tax wages and salary income from the previous year, excluding payments-in-kind or reimbursements.

- **HHINCOME**: Total income of household members aged 15+ from the previous year, expressed in contemporary dollars. Household income includes all members, while family income (FTOTINC) includes only related members.

- **PERWT**: Weights each person to represent the U.S. population for accurate person-level analysis.

- **SERIAL**: A unique household identifier; combines with SAMPLE and PERNUM to uniquely identify each person in the database.

Figure 4.4: Mean hours for four income groups by gender



Note: The income groups (poor, lower middle class, upper middle class, and rich) are defined based on income quartiles for each gender. The mean hours represent the number of hours per week usually worked in the previous year for each group. The variable covers the total population aged between 25 and 55 in each group.

A.4 Additional statistics

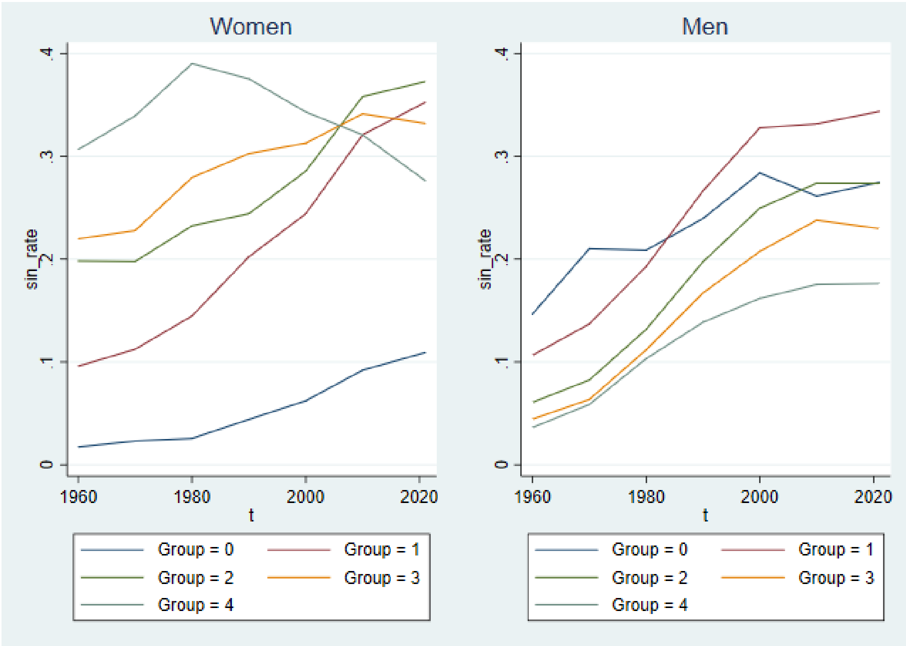
Figure 4.5: Marital rates among 35-44 years old over time by income group.

Note: Group 0 represents non-working individuals. Groups 1, 2, 3, and 4 represent the first, second, third, and fourth quartiles, respectively, by gender. The married fraction represents the ratio of married individuals to the total population aged between 35 and 44 (a la (Chiappori et al., 2020))that is not married. Among all marital categories, only those who are married with a spouse present are considered married. The remaining categories (including married with spouse absent, separated, divorced, widowed, and never married/single) are classified as single.

	husband´s income group			
Wife´s income group	Group 1	Group 2	Group 3	Group 4
	Age cohort 35-44 in 1960			
Group 1	**0.140**	0.063	0.044	0.027
Group 2	0.566	**0.300**	0.214	0.130
Group 3	1	0.600	**0.443**	0.300
Group 4	1.748	0.913	0.701	**0.500**
	Age cohort 35-44 in 2021			
Group 1	**0.649**	0.239	0.138	0.049
Group 2	1.44	**0.745**	0.441	0.235
Group 3	2.765	1.155	**0.758**	0.428
Group 4	6.666	2.075	1.271	**0.795**

Table 4.10: Wife‘s to husband wage gap by income groups and age cohort - Years 1960 and 2021

Figure 4.6: Singlehood rates among 35-44 years old over time by income group.



Note: Group 0 represents non-working individuals. Groups 1, 2, 3, and 4 represent the first, second, third, and fourth quartiles, respectively, by gender. The singles fraction represents the ratio of single individuals to the total population aged between 35 and 44 (a la (Chiappori et al., 2020)).

	husband´s income group				
wife´s income group	Group 0	Group 1	Group 2	Group 3	Group 4
	Age cohort 35-44 in 1960				
Group 0	0.5	9.5	12.7	15.2	19.8
Group 1	0.09	2.8	2.9	3.3	3.5
Group 2	0.1	2.8	3.2	2.9	2.1
Group 3	0.1	2.1	3.0	2.5	1.4
Group 4	0.1	1.1	2.4	2.7	2.1
	Age cohort 35-44 in 2021				
Group 0	0.3	2.9	3.9	4.1	7.04
Group 1	0.5	4.25	4.34	4.35	5.1
Group 2	0.5	3.5	5.5	4.7	3.9
Group 3	0.5	2.5	5.1	6.6	5.3
Group 4	0.7	2.1	3.5	6.4	11.6

Table 4.11: Sorting patterns by income group and Age cohorts - Years 1960 and 2021

Gender	Income Group	Mean
1	1	43.02973247397176
1	2	47.14369023493617
2	1	33.24870552543015
2	2	41.58751660956846

Table 4.12: Average hours worked for married by gender and Income groups - Year 2000

Figure 4.7: Cohabitation rates among 35-44 years old over time by income group.

Note: Group 0 represents non-working individuals. Groups 1, 2, 3, and 4 represent the first, second, third, and fourth quartiles, respectively, by gender. The cohabitation fraction represents the ratio of cohabitating individuals to the total population aged between 35 and 44 (a la (Chiappori et al., 2020)).

Gender	Income Group	Mean
1	1	41.09009027007127
1	2	46.26401800772462
2	1	36.80292590089455
2	2	42.66683717226666

Table 4.13: Average hours worked for singles by gender and Income groups - Year 2000

Gender	Income Group	Mean
1	1	38.89426246884116
1	2	44.92787574329627
2	1	35.61329658168626
2	2	42.42494547195188

Table 4.16: Average hours worked for singles by gender and Income groups - Year 2021

	Income Group 1	Income Group 2
Married Male	0.917	2.324
Married Female	0.594	1.808
Single Male	0.806	2.067
Single Female	0.607	1.653

Table 4.17: Wage Rates by marital status and income group- Year 2021

	Income Group 1	Income Group 2
Married Male	1.000	2.183
Married Female	0.516	1.429
Single Male	0.905	2.024
Single Female	0.556	1.429

Table 4.14: Wage Rates by marital status and income group - Year 2000

Gender	Income Group	Mean
1	1	40.82876601205336
1	2	45.477141640055
2	1	33.48136521995662
2	2	41.99393929212777

Table 4.15: Average hours worked married by gender and Income groups - Year 2021

5 Conclusion

Conclusion

In this thesis, I explore three distinct topics that mutually affect and are impacted by inequality. I analyze how government spending could be detrimental to the income of individuals born in middle-class families, how inequality impacts redistribution through the mechanism of education funding, and lastly, I assess how labor market outcomes affect marriage decisions, which eventually could impact household inequality.

In Chapter 2, using US survey data (ACS) from 2005 to 2009, I document a heterogeneous relationship between government spending on education and the parental decision to opt out of private education. Particularly, I found that middle-income households' decision to opt out of private education is more sensitive to government spending than low and high-income families. Based on this finding, I present an overlapping generation model that addresses the implications of parental decisions on educational investment in a dual education system. By exploring the interaction between government policies, differential education quality, and social class dynamics, I provide insights into how public and private education choices influence inequality and mobility. In specific parametric configurations, government investment can hurt future generations of some middle-income families by instigating opting out of better schools. Future work for this paper will involve using longitudinal data to prove the causality between government spending and negative intergenerational mobility through the latter described mechanism. Additionally, I will study the model at the steady state and simulate it to show how the distribution of income is affected in the long run by policy changes.

In Chapter 3, co-authored with Calin G. Arcalean and Ioana C. Schiopu, we construct a data set using the annual Survey of School System Finances, the American Community Survey – Education Tabulation (ACS-ED) and the National Historical Geographic Information System (NHGIS). We examine the relationship between inequality and redistribution in U.S. school districts and find that rising inequality is associated with a reduction in local public education funding per pupil, particularly in economically disadvantaged districts. These results are maintained after using instrumental synthetic inequality measures based on synthetic income distributions. Future plans for this paper involve strengthening the robustness of our instrumental variables. Particularly, we are now using more data points from the National Historical Geographic Information System (NHGIS) to estimate the midpoint of the open-ended bins by matching estimated Gini coefficients with actual Gini data for each school district.

In Chapter 4, I analyze marriage trends in the U.S. by combining empirical findings with a model that explains shifts in marital patterns. I document new marriage trends in the US from the year 2000 to 2021. I find increased married fractions for rich women and a decrease for all other groups, namely poor and rich men as well as poor women. I propose a general equilibrium model of marriage incorporating wage and work hours differentials to explain these changes. I calibrate the model and use it to assess the impact of the main factors over time. I find that, unlike changes in the wage structure, reducing work hours for men can reproduce the new observed trends. Future work for this paper will be to extend the model in two ways. The first is to endogenize wages and work hours to add more flexibility to the model. The second is to incorporate a work-life balance policy and test how such policies would impact marriage rates and fertility decisions.

Bibliography

Aguiar, M., Bils, M., Charles, K. K., and Hurst, E. (2013). Leisure luxuries and the labor supply of young men. *NBER Working Paper No. 23552.*

Alesina, A., Baqir, R., and Easterly, W. (1999). Public Goods And Ethnic Divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284.

Alonso-Carrera, J., Caballé, J., and Raurich, X. (2012). Fiscal policy, composition of intergenerational transfers, and income distribution. *Journal of Economic Behavior & Organization*, 84(1):62–84.

Anelli, M., Osea, G., and Luca, S. (2021). Robots, marriageable men, family, and fertility. Working paper no. 9378, CESifo.

Arcalean, C. and Schiopu, I. (2015). Inequality, opting-out and public education funding. *Social Choice and Welfare*, 46(4):811–837.

Arcalean, C. and Schiopu, I. (2016). Inequality, opting-out and public education funding. *Social Choice and Welfare*, 46(4):811–837.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms*. *The Quarterly Journal of Economics*, 135(2):645–709.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–1597.

Autor, D. H., Dorn, D., and Hanson, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, 1(2):161–178.

Azimi, E., Friesen, J., and Woodcock, S. (2023). Private schools and student achievement. *Education Finance and Policy*, 18(4):623–653.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political Economy*, 81(4):813–846.

Becker, G. S. (1974). A theory of marriage: Part ii. *Journal of Political Economy*, 82(2):S11–S26.

Becker, G. S. (1994). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. The University of Chicago Press, Chicago, third edition.

Becker, G. S. and Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, 4(3):S1–S39.

Bedi, A. S. and Garg, A. (2000). The effectiveness of private versus public schools: the case of indonesia. *Journal of Development Economics*, 61(2):463–494.

Bénabou, R. (1996). Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance. *American Economic Review*, 86(3):584–609.

Benabou, R. (2000). Unequal societies: income distribution and the social contract. *American Economic Review*, 90:96–129.

Bénabou, R. (2000). Unequal Societies: Income Distribution and the Social Contract. *American Economic Review*, 90(1):96–129.

Benzidia, M., Lubrano, M., and Melindi-Ghidi, P. (2024). Education politics, schooling choice and public school quality: the impact of income polarization. *International Tax and Public Finance*, pages 1–29.

Bettinger, E. P. (2005). The effect of charter schools on charter students and public schools. *Economics of Education Review*, 24(2):133–147.

Binder, A. J. and Bound, J. (2019). The declining labor market prospects of less-educated men. *Journal of Economic Perspectives*, 33(2):163–190.

Blandin, A., John, B., and Fang, Y. (2023). Marriage and work among prime-age men. Working paper no. 23-02 and 2313, Federal Reserve Bank of Richmond and Federal Reserve Bank of Dallas.

Boustan, L., Ferreira, F., Winkler, H., and Zolt, E. M. (2013). Inequality and Local Government: Evidence from U.S. Cities and School Districts, 1970-2000. *The Review of Economics and Statistics*, 95(4):1291–1302.

Brotherhood, L. and Delalibera, B. (2019). The allocation of public expenditures across educational stages: A quantitative analysis for a developing country. Mimeo.

Carnevale, A. P., Fasules, M. L., Quinn, M. C., and Campbell, K. P. (2019). Born to win, schooled to lose: Why equally talented students don't get equal chances to be all they can be? GEW Georgetown Working Paper.

Caucutt, E. M., Guner, N., and Rauh, C. (2021). Is marriage for white people? incarceration, unemployment, and the racial marriage divide. *Journal of Human Resources*, 56(1):153–197.

Cerina, F., Moro, A., and Rendall, M. (2019). The role of gender in employment polarization. *Macroeconomic Dynamics*, 23(8):3475–3488.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states *. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chiappori, P.-A., Iyigun, M., and Weiss, Y. (2018). The market for marriage and divorce. *European Economic Review*, 109:162–175.

Chiappori, P.-A., Salanié, B., and Weiss, Y. (2020). Partner choice, investment in children, and the marital college premium. *American Economic Journal: Applied Economics*, 12(1):76–109.

Chubb, J. E. and Moe, T. M. (1990). *Politics, Markets and American School.* Brookings Institution, Washington, DC.

Church, J. D. (2014). Explaining the shift in consumer expenditures from commodities to services, 1982–2012. *Monthly Labor Review.*

Coleman, J. S., Hoffer, T., and Kilgore, S. (1982). *Secondary School Achievement. Public, Catholic and Private Schools Compared.* Basic Books, Inc., New York.

Corcoran, S. and Evans, W. N. (2009). Income Inequality, the Median Voter, and the Support for Public Education. *Department of Economics, University of Notre Dame.*

Crawfurd, L., Hares, S., and Todd, R. (2023). The impact of private schools, school chains and ppps in developing countries. *The World Bank Research Observer*, 39(1):97–123.

de la Croix, D. and Doepke, M. (2003). Inequality and Growth: Why Differential Fertility Matters. *American Economic Review*, 93(4):1091–1113.

de la Croix, D. and Doepke, M. (2004). Public versus private education when differential fertility matters. *Journal of Development Economics*, 73(2):607–629.

de la Croix, D. and Doepke, M. (2009a). To segregate or to integrate: education politics and democracy. *Review of Economic Studies*, 76(2):597–628.

de la Croix, D. and Doepke, M. (2009b). To Segregate or to Integrate: Education Politics and Democracy. *Review of Economic Studies*, 76(2):597–628.

Doepke, M., Hannusch, A., Kindermann, F., and Tertilt, M. (2022). The economics of fertility: A new era. NBER Working Paper 29948, National Bureau of Economic Research.

Durlauf, S. N., Kourtellos, A., and Tan, C. M. (2022). The great gatsby curve. *Annual Review of Economics*, 14(1):571–605.

Easterly, W. (2001). The middle class consensus and economic development. *Journal of Economic Growth*, 6(4):317–335.

Epple, D. and Romano, R. (1996). Ends against the Middle: Determining Public Service Provision when there are Private Alternatives. *Journal of Public Economics*, 62:297–325.

Ermisch, J., Jäntti, M., and Smeeding, T. (2012). *Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage*. Russell Sage Foundation, New York.

Fan, C.-S., Pestieau, P., and Pang, Y. (2020). Nature versus nurture in social mobility under private and public education systems. Available at SSRN: https://ssrn.com/abstract=3671258 or http://dx.doi.org/10.2139/ssrn.3671258.

Freeman, R. (1992). *How Much Has De-Unionisation Contributed to the Rise in Male Earnings Inequality?*, pages 133–163. Sage Press, NY. NBER WP#3826, 8/91.

Galor, O. and Zeira, J. (1993). Income distribution and macroeconomics. *The Review of Economic Studies*, 60(1):35–52.

Glomm, G. and Ravikumar, B. (1992). Public versus private investment in human capital: Endogenous growth and income inequality. *Journal of Political Economy*, 100:818–834.

Glomm, G. and Ravikumar, B. (1998). Opting out of publicly provided services: A majority voting result. *Social Choice and Welfare*, 15(2):187–199.

Goldin, C. and Katz, L. F. (2008). *The Race between Education and Technology*. Harvard University Press.

Greenwood, J. and Guner, N. (2009). Marriage and divorce since world war ii: Analyzing the role of technological progress on the formation of households. In Acemoglu, D., Rogoff, K., and Woodford, M., editors, *NBER Macroeconomics Annual 2008, Vol. 23*, pages 231–276. University of Chicago Press, Chicago.

Greenwood, J. and Guner, N. (2010). Social change: The sexual revolution. *International Economic Review*, 51(4):893–923.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2014). Marry your like: Assortative mating and income inequality. *American Economic Review*, 104(5):348–353.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2016). Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation. *American Economic Journal: Macroeconomics*, 8(1):1–41.

Hajnal, Z. L. and Lewis, P. G. (2003). Municipal Institutions and Voter Turnout in Local Elections. *Urban Affairs Review*, 38(5):645–668.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24:1141–1177.

Hanushek, E. A. and Woessmann, L. (2011). Chapter 2 - the economics of international differences in educational achievement. volume 3 of *Handbook of the Economics of Education*, pages 89–200. Elsevier.

Harris, A. R., Evans, W. N., and Schwab, R. M. (2001). Education Spending in an Aging America. *Journal of Public Economics*, 81:449–472.

Helpman, E., Itskhoki, O., and Redding, S. (2010). Inequality and unemployment in a global economy. *Econometrica*, 78(4):1239–1283.

Hoyt, W. H. and Lee, K. (1998). Educational vouchers, welfare effects, and voting. *Journal of Public Economics*, 69(2):211–228.

Jerrim, J. and Macmillan, L. (2015). Income inequality, intergenerational mobility, and the great gatsby curve: Is education the key? *Social Forces*, 94(2):505–533. Published: 09 June 2015.

Jimenez, E., Lockheed, M. E., and Wattanawaha, V. P. (1991). The relative efficiency of private and public schools in developing countries. *The World Bank Research Observer*, 6(2):205–218.

Jones, L. E., Manuelli, R. E., and McGrattan, E. R. (2015). Why are married women working so much? *Journal of Demographic Economics*, 81(1):75–114.

Knowles, J. A. (2012). Why are married men working so much? an aggregate analysis of intra-household bargaining and labour supply. *The Review of Economic Studies*, 80(3):1055–1085.

Lindert, P. H. (1996). What Limits Social Spending? *Explorations in Economic History*, 33(1):1–34.

Melindi-Ghidi, P. (2018). Inequality, educational choice, and public school quality in income-mixing communities. *Journal of Public Economic Theory*, 20(6):914–943.

Meltzer, A. H. and Richard, S. F. (1981). A Rational Theory of the Size of Government. *The Journal of Political Economy*, 89(5):914–927.

Mincer, J. A. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research (NBER), New York.

Mizzala, A., Romaguera, P., and Farren, D. (2002). The technical efficiency of schools in chile. *Applied Economics*, 34(12):1533–1552.

Morlan, R. L. (1984). Municipal versus national election voter turnout: Europe and the United States. *Political Science Quarterly*, 99:457–470.

Moro, A., Moslehi, S., and Tanaka, S. (2017). Marriage and economic development in the twentieth century. *Journal of Demographic Economics*, 83(4):379–420.

Neal, D. (1997). The effects of catholic secondary schooling on educational achievement. *Journal of Labor Economics*, 15(1):98–123.

Newman, A. F. (2018). Career women and the durability of marriage. Working Paper, Boston University.

Opdenakker, M.-C. and Van Damme, J. (2006). Differences between secondary schools: A study about school context, group composition, school practice, and school effects with special attention to public and catholic schools and types of schools. *School Effectiveness and School Improvement*, 17(1):87–117.

Piketty, T. and Goldhammer, A. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Piketty, T. and Goldhammer, A. (2020). *Capital and Ideology*. Harvard University Press.

Piketty, T. and Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, 81(5):1851–1886.

Poterba, J. M. (1997). Demographic Structure and the Political Economy of Public Education. *Journal of Policy Analysis and Management*, 16:48–66.

Psacharopoulos, G. (1987). Public versus private schools in developing countries: Evidence from colombia and tanzania. *International Journal of Educational Development*, 7(1):59–67.

Qian, Y. (2017). Gender asymmetry in educational and income assortative marriage. *Family Relations*, 79:318–336.

Regalia, F. and Ríos-Rull, J.-V. (2019). What accounts for the increase in the number of single households? *Review of Economic Dynamics*, 33:1–32.

Restuccia, D. and Urrutia, C. (2004). Intergenerational persistence of earnings: The role of early and college education. *American Economic Review*, 94(5):1354–1378.

Rosenstone, S. J. and Hansen, J. M. (1993). *Mobilization, Participation, and Democracy in America*. Macmillan Publishing Company, New York.

Ruggles, S., Flood, S., Sobek, M., Backman, D., Chen, A., Cooper, G., Richards, S., Rodgers, R., and Schouweiler, M. (2024). IPUMS USA: Version 15.0 [dataset].

Ruggles, S. J., Alexander, T., Genadek, K., Goeken, R., Schroeder, M. B., and Sobek, M. (2010). *Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database].* Minneapolis: University of Minnesota.

Schettino, F. and Khan, H. A. (2020). Income polarization in the usa: What happened to the middle class in the last few decades? *Structural Change and Economic Dynamics*, 53:149–161.

Simula, L. and Trannoy, A. (2010). Optimal income tax under the threat of migration by top-income earners. *Journal of Public Economics*, 94(1):163–173.

Skiba, R., Simmons, A., Ritter, S., Gibb, A., Rausch, M. K., Cuadrado, J., and Chung, C.-G. (2008). Achieving equity in special education: History, status, and current challenges. *Exceptional Children*, 74(3):264–288.

Solon, G. (1992). Intergenerational income mobility in the united states. *The American Economic Review*, 82(3):393–408.

Stevans, L. K. and Sessions, D. N. (2000). Private/public school choice and student performance revisited. *Education Economics*, 8(2):169–184.

Stevenson, B. and Wolfers, J. (2007). Marriage and divorce: Changes and their driving forces. *Journal of Economic Perspectives*, 21(2):27–52.

Stiglitz, J. E. (2012). *The Price of Inequality: How Today's Divided Society Endangers Our Future.* W.W. Norton Co., New York.

Verba, S., Lehman, S. K., and Brady, H. E. (1995). *Voice and Equality: Civic Voluntarism in American Politics.* Harvard University Press, Cambridge, Mass.

Títol: Assaigs sobre l'educació, el matrimoni i les classes socials.

Resum

Aquesta tesi investiga els mecanismes a través dels quals es crea i es perpetua la desigualtat econòmica. Basant-se en diverses teories econòmiques i evidència empírica, examina tres temes interconnectats que contribueixen a la persistència de la desigualtat.

El primer capítol explora l'impacte de la despesa governamental en educació, revelant una paradoxa on l'augment de la inversió pública en educació pot desavantatjar desproporcionadament les llars de renda mitjana en influir en les seves decisions d'optar per l'escolarització privada. Utilitzant un model de generacions superposades, destaca les implicacions d'aquesta dinàmica en la mobilitat intergeneracional i les disparitats d'ingressos.

El segon capítol analitza la relació entre la desigualtat d'ingressos i el finançament local de l'educació pública als districtes escolars dels Estats Units, demostrant que una major desigualtat condueix a una reducció del finançament local, particularment en zones econòmicament desafavorides. S'empra una estratègia de variables instrumentals per abordar problemes d'endogeneïtat. Les troballes emfatitzen la necessitat de polítiques educatives equitatives que mitiguen aquests efectes adversos.

El capítol final investiga les tendències canviants del matrimoni als Estats Units, centrant-se en el paper de les estructures salarials i les hores de treball en la configuració de les dinàmiques familiars i la desigualtat. Es desenvolupa un model d'equilibri general per analitzar la interacció de factors econòmics i temporals en les decisions matrimonials. Les troballes indiquen que els canvis en les hores de treball juguen un paper crucial en l'explicació de les tendències recents del matrimoni.

En conjunt, aquesta dissertació contribueix a comprendre com l'educació, la redistribució i el matrimoni interactuen per sostenir la desigualtat i proposa vies per a futures investigacions i intervencions polítiques destinades a mitigar la desigualtat i fomentar la mobilitat.

Paraules clau: Finançament de l'educació, Desigualtat, Transferències intergeneracionals, Matrimoni, Economia política, Educació pública i privada.