

Hennig, Jan-Luca

**Conference Paper**

## Labor Market Polarization and Intergenerational Mobility: Theory and Evidence

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Hennig, Jan-Luca (2021) : Labor Market Polarization and Intergenerational Mobility: Theory and Evidence, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242353>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Labor Market Polarization and Intergenerational Mobility: Theory and Evidence

Jan-Luca Hennig\*  
Trinity College Dublin

February 5, 2021  
Latest Version

## Abstract

*In this paper I investigate the causal relationship between labor market polarization and intergenerational mobility, two of the most important features of advanced labor markets in recent decades. The former relates to the disappearance of middle-wage routine jobs and the rise of both high- and low-income jobs. The latter measures the cross-generational link between the income of parents and that of their children. To understand the implications of polarization on economic mobility, I develop a model that accounts for labor market polarization, spatial heterogeneity and intergenerational mobility. Specifically, I incorporate the task framework into an overlapping generations model, which models education choice as a function of parental bequests and future wages. The main predictions from the model encompass limited upward mobility for children from low-income parents if replacement of routine labor is high, and educational polarization for younger generations. I test the model predictions exploiting various data sources for the United States. Empirical analyses provide evidence for lower upward mobility for children from low-income parents, higher intergenerational elasticity for children whose parents work in manual occupations, and for educational polarization of young labor market entrants. (JEL J21, J23, J24, J62, O33, R23)*

**Keywords:** Labor Market Polarization, Intergenerational Mobility, Educational Choice, Local Labor Markets

## 1 Introduction

The “American Dream” allows everyone to be successful regardless of their geographic and family background. However, the United States has turned into one of the least socially

---

\*Trinity College Dublin (E-mail: [hennigj@tcd.ie](mailto:hennigj@tcd.ie)). I am grateful to my thesis supervisors Fadi Hassan and Davide Romelli for all of the comments and remarks aiding me to advance this paper. I would like to thank all members of the Economics Department of Trinity College for helpful comments, specifically to Joseph Kopecky and Michael Wycherly. I further want to thank Donald Davis, Réka Juhász, Sandra Black and participants of the International Economics Colloquium at Columbia University. I also want to thank Leah Platt Boustand and David Dorn for helpful comments. I am also grateful for questions from participants at the EEA Virtual 2020 and the IAB Conference on Labor Market Transitions. Parts of this project were completed while at Columbia University, whose hospitality is gratefully appreciated. For financial support I want to thank the Grattan Scholars program.

mobile countries among advanced economies, and with stark differences within the United States. Chetty et al. (2014) provide evidence of large differences in upward mobility across US commuting-zones ranging from the most mobile to the least mobile among developed countries. Since the 1980s, many advanced economies including the United States have experienced strong labor market polarization. It is defined as the decline in routine occupations and the simultaneous rise of both low-income manual and high-income abstract employment. Autor and Dorn (2013) show that the decline in routine employment also depends on local factors.

This paper establishes a causal relationship between labor market polarization and intergenerational mobility. Labor market polarization can lower equality of opportunity for children from a low-income background via two channels, namely labor supply and labor demand. The first mechanism relates to educational choice, where the highest level of education yields the largest returns, i.e. the wage of an abstract job, but also incurs the highest institutional cost. Children from low-income families are less able to finance college tuition or they are more reluctant to take up a loan to pay for college education. These limited financial resources imply that they opt for either routine or manual jobs. The former also incurs an institutional cost, and considering declining wages (together with falling employment) of routine occupations, obtaining secondary education for taking up routine employment reduces lifetime utility. Therefore, labor market polarization influences education choices of younger generations.

The second channel refers to labor demand, i.e. how many jobs in routine occupations will be lost due to the acquisition of computer capital. Jaimovich and Siu (2012) implicitly provide evidence for the labor demand channel by showing that in all recessions since 1991 jobless recoveries are strongly related to the decline in routine employment. Along the same line, Hershbein and Kahn (2018) use vacancy-postings and show that the demand for routine skills falls and confirms with the interpretation of firms restructuring of production toward routine-biased technologies and higher-skilled workers. The influence of business cycle fluctuations on routine employment and the evidence from vacancy postings is in line with a changes in changing demand for task types.

To fully understand the implications of labor market polarization on equality of opportunity across the income distribution, I set up a simple model drawing on existing models for educational choice and labor market polarization. The model builds on the standard features of the task framework from Autor et al. (2003) and more formalized in Autor and Dorn (2013). The final tradable good is produced with three types of labor, namely manual, routine and abstract, and it features substitutability between routine labor and capital. Over time, labor market polarization is driven by an exogenous decline in the price for capital as in Davis et al. (2020). In order to capture spatial patterns, I assume exogenous local productivity differences, which allows me to further characterize regions in terms of population size, and how locations differ in their degree of labor market polarization.

Households are characterized by an overlapping-generations (OLG) model with educational choice similar to Maoz and Moav (1999). I assume that individuals are upward mobile if they

choose a higher level of education than their parents as returns to education rise with educational levels. Congruent with the task framework, I assume three levels of educational attainment, which differ in their costs. Individuals face educational frictions to attain education, such as ability. The education choice of depends on parental bequests and future wages in each sector. The latter is crucial to understand education decisions in the face of labor market polarization, which reduces routine wages, and therefore returns to secondary education. Bequests is particularly important as Keane and Wolpin (2001) highlight the importance of parental transfers for the young generation to finance educational attainment. Further, they help to understand which increasingly individuals choose primary and tertiary education.

The model delivers multiple testable predictions with respect to intergenerational mobility across time and space via educational choices. As the price for capital, and hence routine wages, fall over time, the model predicts less occupational cross-generational transitions for children whose parents work in either manual or abstract occupations, and more transitions in both directions for children whose parents work in routine occupations in all locations. This translates into less upward mobility for children from parents in manual occupations. High productivity pushes up wages in all occupations, therefore population is larger in more productive locations, which in turn raises housing prices until real wages are equalized across space. Consistently, routine employment is higher in more productive regions, creating a stronger incentive to destroy routine jobs. Therefore, high-productivity locations demand more capital, and experience stronger labor market polarization, and therefore lower upward mobility. As upward mobility is defined as a cross-generational transition if a child acquires a higher level of education than their parent, these model predictions imply that labor market entrants increasingly choose “polar” educational attainment.

In the next step, I confront the main model predictions with the data. First, I test the key prediction that labor market polarization reduces intergenerational mobility across space. Exploiting commuting-zone variation, I identify the causal impact of labor market polarization on intergenerational mobility using an instrumental variable specification, which allows me to circumvent various issues of endogeneity. Based on the model and previous findings by Autor and Dorn (2013), I use historical density and routine intensity as instruments for labor market polarization. I take data on intergenerational mobility from Chetty et al. (2014) and use absolute upward mobility as a benchmark measure because it concentrates on outcomes for children from the 25<sup>th</sup> percentile of the parental income distribution. In line with theoretical predictions, empirical evidence suggests that labor market polarization significantly reduces intergenerational mobility for children from low-income parents.

To test whether the key model prediction holds over time, I estimate the intergenerational elasticity (IGE) depending on parental occupational background using the Panel Study of Income Dynamics (PSID). Specifically, I estimate IGE for children whose parents work in manual, routine or abstract occupations based on the definitions by Autor and Dorn (2013). The results confirm the model predictions, namely a higher IGE for children with parents in manual and ab-

stract occupations, while it is lower when parents work in routine occupations. This is because children with parents in routine jobs, i.e. the middle class, experience more cross-generational transitions and are more probable to enter either manual or abstract employment. Importantly, this pattern develops over time, and first emerges after the investment boom in capital during the 1990s. These findings are in line with “stickiness” among the richest and poorest in society as found by Blanden et al. (2004) for Britain and the U-shaped pattern of intergenerational mobility for the US found by Palomino et al. (2018).

In the last step of the empirical exercise, I confirm that educational attainment is polarizing over time, that it is geographically linked to labor market polarization and depends on family background. I start by showing that education became more polarized in the United States between 1970 and 2018 among labor force participants between 20 and 29 for the United States as a whole. The major increase in educational polarization occurs between 1990 and 2000, which also coincides with the investment boom in capital in the form of information and communication technology. Next, I use the PSID to compute family premia based on Checchi et al. (2013) and polarization indices for various degrees based on parental backgrounds. Third, I exploit variation across time and commuting-zones to show that labor market polarization increases educational polarization. Last, using the Current Population Survey (CPS), I estimate a linear probability model, which provides evidence that educational stickiness is stronger if labor market polarization increases.

The results indicate that a one percentage point increase in labor market polarization decreases the expected rank of children whose parents are at the 25<sup>th</sup> percentile of the income distribution by .57 ranks in the national income distribution. This is a sizeable economic effect of labor market polarization on intergenerational mobility as children in commuting zones without labor market polarization would be ranked 4.42 percentiles higher than children from the median-hit commuting zone. This roughly translates into 13.79% higher incomes, which is equivalent to nearly 3600 USD.<sup>1</sup>

Labor market polarization of wages and employment driven by automation is closely linked to income inequality as argued by Acemoglu and Restrepo (2020). Krueger (2012) and Corak (2013) highlight the negative link between income inequality and intergenerational mobility, a relationship known as the Great Gatsby curve. It holds across countries and within the United States over time. This relationship is relevant because it implies that higher inequality, usually measured with the Gini coefficient, magnifies the “persistence in the advantages and disadvantages of income passed from parents to the children”. At the same time, stronger labor market polarization is also associated with widening wage inequality. However, the underlying mechanisms how income inequality impacts upward mobility are opaque.

Linking labor market polarization to intergenerational mobility has two major advantages over the link between income inequality and social mobility. First, the Gini coefficient is an annual snapshot of inequality, when, as noted already by Sahota (1978), most policy-advice to reduce

---

<sup>1</sup>See Chetty and Hendren (2018b).

income disparities is concerned with lifetime inequality. Huggett et al. (2011) find that initial conditions (as of age 23) account for 61.5% of lifetime earnings, confirming that policy should focus on affecting early decisions to accumulate human capital. Labor market polarization, compared to inequality as an outcome variable, directly affects the decision of human capital investment towards either high-or low-skill education, which implies higher lifetime inequality. In the same vein, Dabla-Norris et al. (2019) show that especially young workers have to adjust to the evolution of labor markets.

The second advantage of using labor market polarization over income inequality refers to the differences in quality of jobs associated with low- and high-skilled employment. The quality of jobs encompasses a multitude of different characteristics, Osterman and Shulman (2011) argue that quality of jobs is also polarized, with “good” jobs providing anywhere from living to astronomically high wages, benefits, opportunities for advancement and training, and “bad”, dead-end jobs without career-progression possibilities and paying minimum or near-minimum wages. They claim that about 25% of all jobs in the United States are “bad” jobs. Major et al. (2019) also mention the changing nature of jobs, i.e. they refer to jobs in the gig economy, which has created millions of manual jobs often done by the solo self-employed, lacking security, progression or rights. Harrison and Bluestone (1990) argue that the introduction of new management practices, which focuses on making labor, in particular low-skill labor, a more variable factor of production fostered wage polarization in the United States.

A large recent literature focuses on various explanations of intergenerational mobility, in particular sparked by Chetty et al. (2014). The authors sort factors which have been associated with intergenerational mobility in the literature into various categories. The availability of intergenerational mobility within the United States provided by the Equality of Opportunity Projects sparked a surge in estimating causal impacts of factors associated with upward mobility. A non-exhaustive list of this research includes Chetty and Hendren (2018a) and Chetty and Hendren (2018b), who investigate the impact of neighborhoods on children outcomes and find strong childhood exposure effects. Gallagher et al. (2019) identify the importance of family structure. Sharkey and Torrats-Espinosa (2017) examine how exposure to violent crime reduces intergenerational mobility. Andrews et al. (2017) show that areas with higher historical racial segregation exhibit lower levels of upward mobility.

This study contributes to the rising literature investigating the determinants of intergenerational mobility by investigating the role of labor market polarization and automation. Rothstein (2019) focuses on the quality of schools and their effect on intergenerational mobility, but finds little evidence that school quality plays a significant role. He concludes that the structure of local labor markets is a likely factor influencing economic mobility.<sup>2</sup> Tan (2019) provides evidence that labor markets matter historically, in particular industrial job opportunities were important drivers of upward mobility in the early 20<sup>th</sup> century.

This paper is closest to Adão et al. (2020), who focus on two margins of adjustment to tech-

---

<sup>2</sup>He also mentions job networks and marriage markets as other determinants of intergenerational mobility.

nological progress, namely a within-generation reallocation of labor and cross-generational adjustments in the skill distribution. They provide evidence that the recent trend of innovations towards cognitive occupations affected the young generation stronger. They highlight that technological transitions are slower and more unequal if innovations are directed towards skills which are not abundant in the contemporaneous labor force. The main difference between the two papers is that I focus on the impact of the young generation across the parental income and skill distribution, and analyze how labor market polarization - driven by automation - impacts various labor market outcomes of children from different parental backgrounds, specifically educational choice and intergenerational mobility.

In the next section, I outline the theoretical framework combining existing models of educational choice and labor market polarization. Section 3 presents all data sources used in the empirical analysis, in particular the Decennial Censuses, the Panel Study of Income Dynamics, and the data on intergenerational mobility by Chetty et al. (2014). Section 4 takes the model predictions to the data. Section 5 concludes.

## 2 Theoretical framework

In this section, I build a model incorporating the task framework as in Autor and Dorn (2013) into an overlapping-generations model where individuals choose their education as in Maoz and Moav (1999). The model features the substitutability between routine labor and capital, and exogenous location-specific productivity differences. Individuals choose between three levels of educational attainment, and their education decision depends on parental bequests and future wage ratios. Primary education allows individuals to work in manual employment, secondary education in routine employment, and tertiary education in abstract employment.

The key predictions of the model relate to how labor market polarization affects educational choice, and therefore intergenerational mobility, across time and space. First, as the price for capital is declining exogenously over time, individuals from all backgrounds are less likely to choose secondary education. Therefore, the share of routine employment declines over time, while the employment shares of manual and abstract labor increase. The second key prediction applies to regional variation in labor market polarization due to idiosyncratic demand for capital. Exogenously more productive locations demand more capital, and therefore labor market polarization is stronger. As a consequence, the adverse impact on children from low-income parents is stronger in more productive regions.

### 2.1 Production technology

The economy produces a final good in every period  $t$  combining labor of three different skill-levels and capital. The good is tradable at no costs across locations, and is produced with the

following production technology:

$$A_j F(L_{mjt}, L_{rjt}, L_{ajt}, K_{jt}) = A_j \left[ A_m L_{mjt}^{\gamma_m} + (A_r L_{rjt}^\theta + A_k K_{jt}^\theta)^{\frac{\gamma_r}{\theta}} + A_a L_{ajt}^{\gamma_a} \right], \quad (1)$$

where  $A_j$  denotes total productivity in location  $j$ , and  $A_i$  represents factor-augmenting technology for employment type  $i \in \{a, m, r\}$ , where  $a$  stands for abstract (high-skill),  $m$  for manual (low-skill) and  $r$  for routine (middle-skill).  $L_{ijt}$  indicates employment of type  $i$  in region  $j$  at time  $t$ . The fourth factor of production is capital  $K_{jt}$ , which also provides routine tasks. To assure substitutability between routine labor and capital, it must hold that  $\gamma_r < \theta$ . Capital often refers to information and communication technology (ICT), and encompasses both hardware and software.

The production of capital ( $K_t$ ) is analogous to Davis et al. (2020), and is produced using the following technology:

$$K_t = \frac{1}{\zeta_t} Q_t, \quad (2)$$

where  $Q_t$  is the amount of final goods and  $\zeta_t$  is a technology parameter. Perfect competition for the intermediate good implies:

$$p_{kt} = \zeta_t \quad (3)$$

The technology parameter, and hence the price of ICT capital, declines exogenously over time. Unsurprisingly, a lower price for capital induces a larger capital stock in the economy, as  $\frac{\partial K_t}{\partial \zeta_t} < 0$ . For simplicity I assume that capital as intermediate good fully depreciates every period.

As Davis et al. (2020) explain, the intermediate good  $K_t$  has two interpretations: first, it is a capital good that substitutes for middle-paid labor as in Autor and Dorn (2013). With this view,  $\zeta_t$  is a parameter that governs the efficiency of producing the capital good. The second interpretation is that  $K_t$  is an imported intermediate and  $\zeta_t$  then denotes the terms of trade. As a result, a drop in  $p_{kt}$  could be either due to routinization, i.e. a drop in the price of computer capital, or due to offshoring, i.e. a drop in the domestic price of the intermediate import due to technical progress abroad or the removal of trade barriers.

All firms are price-takers and do not affect wages. Wages are determined simultaneously for all skills  $i$  and all locations  $j$ . The price of capital is governed by the exogenous technology parameter  $\zeta_t$ . The firm's profit maximization problem takes on the following form:

$$\max_{L_{ijt} \forall i} A_j F(L_{mjt}, L_{rjt}, L_{ajt}, K_{jt}) - \sum_{L_{ijt} \forall i} w_{ijt} L_{ijt} - p_{kt} k_{jt}, \quad (4)$$

where the first-order conditions (FOCs) of each labor type and capital are equal to the respective wage and the price of capital, respectively. The constraint is that each labor type and capital are greater or equal to zero, i.e.  $L_{ijt} \geq 0 \forall i$  and  $K_{jt} \geq 0$ .

I derive the FOCs for each employment type ( $L_{ijt}$ ) and capital ( $K_{jt}$ ), which hold in each location



$j$ :

$$w_{mjt} = A_j \gamma_m A_m L_{mjt}^{\gamma_m - 1} \quad \forall j \quad (5a)$$

$$w_{rjt} = A_j \gamma_r (A_r L_{rjt}^\theta + A_k K_{jt}^\theta)^{\frac{\gamma_r - \theta}{\theta}} A_r L_{rjt}^{\theta - 1} \quad \forall j \quad (5b)$$

$$w_{ajt} = A_j \gamma_a A_a L_{ajt}^{\gamma_a - 1} \quad \forall j \quad (5c)$$

$$p_{kt} = A_j \gamma_r (A_r L_{rjt}^\theta + A_k K_{jt}^\theta)^{\frac{\gamma_r - \theta}{\theta}} A_k K_{jt}^{\theta - 1} \quad \forall j. \quad (5d)$$

The FOCs in equations (5a) to (5c) reveal that wages of each skill type  $i$  depend on four factors. First, on location-specific exogenous productivity  $A_j$ . Second, on the output elasticity of each labor type  $\gamma_i$ . Third, the skill-specific productivity parameter  $A_i$ , and fourth, on labor input of the specific skill  $L_{ijt}$ . Additionally, the wage of routine workers also depends on the capital stock. Due to the substitutability between routine labor and ICT capital ( $\gamma_r < \theta$ ), a rising capital stock decreases returns to routine tasks.

Manual and abstract wages are unaffected by falling capital prices. While the substitutability between routine labor and capital is common to the literature on labor market polarization, it is more divided on how ICT capital affects the wages of extreme skills. Autor and Dorn (2013) assume complementarity between computer capital and abstract labor, Davis et al. (2020) indicates complementarity of capital with both high and low-skilled workers. vom Lehn (2019) allows for the most flexible specification with three nests in the CES production function. He argues that ICT capital is complementary to abstract labor, and that the relationship of manual employment with other employment types is more ambiguous. For simplicity, I follow Eeckhout et al. (2019) and assume that ICT capital has no impact on manual and abstract wages.

## 2.2 Education decision

Each household consists of one parent and one child, and each individual lives for two periods. Individuals gain utility from consumption in the first period, and from consumption, housing and bequests to their children in the second period.<sup>3</sup> Workers of all types are perfectly mobile across regions without migration costs, implying utility equalization for a given type across locations. All individuals have the same preferences according to the quasi-linear utility function:

$$U_{ij} = \log U_{ijt} + \log U_{ijt+1}, \text{ where} \quad (6)$$

$$U_{ijt} = c_{ijt}, \text{ and} \quad (7)$$

$$U_{ijt+1} = c_{ijt+1}^\alpha h_{ijt+1}^\beta x_{ijt+1}^{1-\alpha-\beta} \quad (8)$$

---

<sup>3</sup>I am using the terms “bequest” and “parental investment” synonymously. This is because the former is used in the previous literature, but it can also be interpreted as the latter because education takes place during the first period in life, hence parents need to invest the money at the beginning of the first period, not at its end.

where  $c_{ijt}$  and  $c_{ijt+1}$  denote consumption of an individual  $i$  born in period  $t$  and living in region  $j$  in the two periods of her life ( $t$  and  $t + 1$ ),  $h_{ijt+1}$  represents housing costs and  $x_{ijt+1}$  illustrates the transfer to her child. The bequest motive does not differ between the different types of individuals. Relaxing this assumption would reinforce the results of the model.<sup>4</sup>

In the first period, individuals receive parental bequests from which they finance current consumption and their education. Similar to Maoz and Moav (1999) and Owen and Weil (1998), financial markets where individuals can borrow to finance higher levels of education are absent. This is in line with the argumentation by Tobin (1982) that lenders do not accept future human capital as collateral. In the second period, households obtain a wage according to the education they obtained in the first period and divide their labor income between consumption, housing and bequest to their children. Hence, households face the following budget constraints:

$$x_{ijt} = c_{ijt} + \delta_i \tau_{ijt} \quad (9a)$$

$$w_{ijt+1} = c_{ijt+1} + p_{jt+1} h_{ijt+1} + x_{ijt+1}, \quad (9b)$$

where  $\delta_i$  symbolizes institutional costs for the three different types of education, and  $\tau_{ijt}$  indicates individual educational frictions.<sup>5</sup> Hence, total education costs depend on institutional costs and individual educational frictions.

Individual educational frictions raise the costs of educational frictions and are paid in the form of first-period consumption. They can be thought of as a function of various factors, e.g. inverse ability as in Maoz and Moav (1999) and Abraham (2008), informational constraints, family support and engagement as highlighted by Mayer et al. (2019), preferences, or access to funding.<sup>6</sup> For simplicity, and in line with Papageorge and Thom (2020), individual educational frictions are independent from parental background. Relaxing this assumption and allowing for a positive relationship between parental investment (reflecting parental income) and individual educational frictions would reinforce the results obtained below. The distribution of educational frictions is constant across locations and time. For simplicity, I assume that educational frictions  $\tau_{ijt}$  are uniformly distributed (within and across cities) in the interval  $(\underline{\tau}_{ijt}, \bar{\tau}_{ijt})$ , where  $\underline{\tau}_{ijt}$  equals minimum educational frictions, and  $\bar{\tau}_{ijt}$  indicates maximum education frictions. The higher educational frictions, the stronger the payment towards educational attainment in the form of consumption in the first period.

---

<sup>4</sup>Empirical evidence that high-income parents have a higher bequest motive is shown by e.g. Menchik and David (1983), and I show below that parents with higher incomes provide more support to their children.

<sup>5</sup>Usually, models considering intergenerational mobility and educational choice consider only two types of education, namely unskilled vs. skilled. One notable exception is Fender (2005), who extends the standard framework by considering self-employment and by introducing a third period, giving agents the chance to choose a sector of work over two periods of their lives.

<sup>6</sup>Albeit I abstract from this issue in the model, educational frictions can also be driven by lower access to higher education as highlighted by Hillman (2016).

Institutional costs take on the following form:

$$\delta_i = \begin{cases} 0 & \text{individual acquires primary education} \\ 1 & \text{individual acquires secondary education} \\ z & \text{individual acquires tertiary education,} \end{cases}$$

where I assume that primary education does not incur costs, which holds in advanced economies as compulsory basic education is provided freely and equally to all children. Let  $z > 1$  denote the time-invariant cost of tertiary education relative to secondary education. Importantly, there are no differences in educational quality by region.

There is no uncertainty in receiving the according wage of household  $i$  in the second period of her life, for which the individual gets an education during the first period. For notational purposes, it is important to mention that the level of educational frictions ( $\tau_{ijt}$ ) is the only parameter in the model, which varies across individuals. As soon as an individual has sorted herself into one level of education, and implicitly occupational group, workers are identical within each group. Therefore, I denote  $i$  for each group instead of for each individual household. Future labor earnings are:

$$w_{ijt+1} = \begin{cases} w_{mjt+1} & \text{if } \delta_i = 0 \\ w_{rjt+1} & \text{if } \delta_i = 1 \\ w_{ajt+1} & \text{if } \delta_i = z, \end{cases}$$

where, as before,  $m, r$  and  $a$  represent the different types of labor in production, namely manual, routine and abstract, respectively.

Since the utility function is separable and there are no capital markets, the individual's maximization of utility can be done backwards in two stages. First, the individual decides on how to allocate her labor income in the second period between consumption, housing and bequest, and then the individual decides on the level of human capital in the first period. The second period maximization implies solving:

$$z(w_{ijt+1}) \equiv \max(c_{ijt+1}^\alpha h_{ijt+1}^\beta x_{ijt+1}^{1-\alpha-\beta}) \quad \text{s.t. (9b)}$$

The maximization of the second period yields the respective equilibrium allocation of resources, with  $c_{ijt+1}^* = \alpha w_{ijt+1}$ ,  $x_{ijt+1}^* = (1 - \alpha - \beta)w_{ijt+1}$  and  $h_{ijt+1}^* = \frac{\beta w_{ijt+1}}{p_{jt+1}}$ . Plugging these results back into  $z(w_{ijt+1})$  gives the indirect utility for the second period. Due to perfect labor mobility, utility is constant across locations in every period. In what follows, for notational simplicity I will compare two cities instead of a number of  $J$  cities, but the results hold without loss of generality. Using the result for optimal housing expenditure ( $h_{ijt+1}^*$ ), I can show that wage ratios across cities relate:

$$\frac{w_{i1t}}{p_{1t}^\beta} = \frac{w_{i2t}}{p_{2t}^\beta} \quad \forall \quad i \in (m, r, a). \quad (10)$$

Equation (10) indicates that real wages for a given occupational group are constant across cities. This finding and the assumption of no migration costs allow me to concentrate on educational choice of individuals *within* a location and to concentrate on nominal income differences between occupations.

Individuals face two simultaneous choices about educational choice, namely whether to choose primary, secondary or tertiary education. To determine the “marginal” individual, i.e. the individual which is indifferent between choosing between two levels of educational attainment, I always compare lifetime utility for two education levels. This yields two thresholds of educational frictions for each parental background, and hence I obtain six six thresholds of educational frictions in total. I use the optimal allocation of resources in the second period and plug them into the utility functions from equation (6) and exploit equation (9a) in order to determine the educational choice. To illustrate the choice problems individuals face, I illustrate two choice problems. Children whose parents work in a manual job, and who choose between primary and secondary education. She will choose secondary education, i.e. routine employment, if and only if:

$$\log(x_{mjt} - \tau_{ijt}) + \log w_{rjt+1} \geq \log x_{mjt} + \log w_{mjt+1}.$$

Analogously, a child whose parents in routine employment, and who chooses between secondary and tertiary education. She will choose tertiary education if and only if:

$$\log(x_{rjt} - z\tau_{ijt}) + \log w_{ajt+1} \geq \log(x_{rjt} - \tau_{ijt}) + \log w_{rjt+1}.$$

From these two examples of the six choice problems, it follows that she will invest in the higher level of educational attainment if educational frictions  $\tau_{ijt}$  are small enough, i.e. if educational costs are not too high. Let  $\hat{\tau}_{ijt}$  denote the critical value of educational frictions for the marginal individual  $i$ , i.e. the individual indifferent between the two choices. The six thresholds take on the following form:

$$\hat{\tau}_{ijt}^{mr} = x_{mjt} \left[ 1 - \frac{w_{mjt+1}}{w_{rjt+1}} \right], \quad (11)$$

$$\hat{\tau}_{ijt}^{ma} = \frac{x_{mjt}}{z} \left[ 1 - \frac{w_{mjt+1}}{w_{ajt+1}} \right] \quad (12)$$

$$\hat{\tau}_{ijt}^{rm} = x_{rjt} \left[ 1 - \frac{w_{mjt+1}}{w_{rjt+1}} \right], \quad (13)$$

$$\hat{\tau}_{ijt}^{ra} = \frac{x_{rjt} \left[ 1 - \frac{w_{rjt+1}}{w_{ajt+1}} \right]}{z - \frac{w_{rjt+1}}{w_{ajt+1}}} \quad (14)$$

$$\hat{\tau}_{ijt}^{am} = \frac{x_{ajt}}{z} \left[ 1 - \frac{w_{mjt+1}}{w_{ajt+1}} \right] \quad (15)$$

$$\hat{\tau}_{ijt}^{ar} = \frac{x_{ajt} \left[ 1 - \frac{w_{rjt+1}}{w_{ajt+1}} \right]}{z - \frac{w_{rjt+1}}{w_{ajt+1}}}, \quad (16)$$

where the first letter of the superscript indicates the occupation in which the individual’s parents worked, and the second letter where the individual works in the second period. For example, in equation (11), the individual’s parents worked in manual employment, and the individual chooses routine employment if her level of educational frictions is below the threshold value  $\hat{\tau}_{ijt}^{mr}$ . I define “upward mobility” with respect to education - and hence implicitly for oc-

cupation and income - if the individual chooses a higher level of education than her parents. Downward mobility is defined as the case where the individual chooses a lower level of educational attainment than her parents.

All thresholds of individual educational frictions are expressed as a function of future wages and of the transfer received individual received from her parents. The components of the threshold functions are similar to Maoz and Moav (1999), Owen and Weil (1998) and Galor and Tsiddon (1997). The inclusion of future wages implies perfect foresight of how wages develop in the future with declining price for capital. Parental transfers play a crucial role in educational choice. This finds empirical support by Keane and Wolpin (2001), and I also provide evidence for stronger financial support by more affluent parents in the empirical section.

I conduct a comparative statics analysis in order to understand how parental bequests and future wages affect the thresholds of educational frictions. Note that the threshold is continuous and differentiable with respect to parental investment  $x_{ijt}$ , and future wages. Consider  $w_{ijt+1}$  the wage for the employment type which requires a relative lower level of educational attainment (numerator), and  $w_{i'jt+1}$  the wage for the employment type which requires a higher level of educational attainment (denominator). The comparative statics reveal the following:  $\frac{\partial \hat{\tau}_{ijt}}{\partial x_{ijt}} > 0$ ,  $\frac{\partial \hat{\tau}_{ijt}}{\partial w_{ijt+1}} < 0$  and  $\frac{\partial \hat{\tau}_{ijt}}{\partial w_{i'jt+1}} > 0$ . In words, these findings imply that higher bequests and a rising wage in  $i'$  raise the threshold of educational frictions, implying that the marginal individual faces higher constraints. This, in turn, means that upward mobility is more likely to occur as more individuals will enter the sector requiring a higher level of education. On the other hand, if wages in  $i$  are rising, then the marginal individual has a lower level of educational frictions, and thus less upward mobility takes place.

When tertiary education, i.e. working in abstract employment, constitutes one side of the choice problem, then the threshold function of individual educational frictions also includes the costs of tertiary education  $z$ . The derivative sign of each threshold including these costs is negative with respect to  $z$ , i.e.  $\frac{\partial \hat{\tau}_{ijt}}{\partial z} < 0$ . This comparative static indicates that the marginal individual has a lower level of educational frictions  $\tau_{ijt}$  if the costs of tertiary education are rising. In other words, upward mobility (from parents with either manual or routine employment) is more difficult if the cost of tertiary education is high.

### 2.3 Upward and Downward Mobility

The model predicts how intergenerational mobility changes due to rising automation of routine tasks varies in two dimensions, namely across time and space. The first dimension emerges because the technology parameter  $\zeta_t$  falls exogenously over time. The declining price for ICT capital implies a rising capital stock in the economy. As ICT capital performs routine tasks, this development has a detrimental impact on routine wages. Therefore, the returns to secondary education fall, while a higher capital stock does not directly affect the returns to primary nor tertiary education, i.e. manual and abstract employment, respectively. In the choice problem of

educational attainment, the wage ratios between different task types matters and influences the level of individual educational frictions.

**Proposition 1. Polarization over Time** Assume substitutability between routine labor and capital, i.e.  $\gamma_r < \theta$ . Substitutability implies  $\frac{\partial w_{rjt}}{\partial K_{jt}} < 0 \forall j$ . Then educational choice of individuals polarizes over time as the capital stock increases because  $\frac{\partial \tau_{ijt}^{mr}}{\partial w_{rjt}} > 0$  and  $\frac{\partial \tau_{ijt}^{rm}}{\partial w_{rjt}} > 0$ , and  $\frac{\partial \tau_{ijt}^{ra}}{\partial w_{rjt}} < 0$  and  $\frac{\partial \tau_{ijt}^{ar}}{\partial w_{rjt}} < 0 \forall j$ . The first two derivatives indicate that falling routine wages induces more individuals whose parents work in either manual or routine employment acquire primary education instead of secondary education. The latter two derivatives indicate that falling routine wages induces more individuals whose parents work in either routine or abstract employment acquire tertiary education instead of secondary education.

*Proof see Appendix A.1.*

The second dimension concerns difference in cross-generational transitions in education, and therefore occupational type and income, across space. The underlying idea is to determine whether demand for capital is location-specific. Locations only differ by their exogenous productivity level  $A_j$ , i.e. some locations have a higher absolute advantage across all employment types. As argued before, the locations-specific productivity parameter is a determinant of wages for all skill types. Before turning to the actual prediction of interest, I derive a key property as to how locations differ due to differences in absolute advantage in production of the final good, which I will also exploit in my empirical analysis. As noted above, I will compare two cities instead of a number of  $J$  cities without a loss of generality. In what follows, I always assume that productivity in location 1 is higher than in location 2, i.e.  $A_1 > A_2$ .

This property concerns the population size (and density) of different locations. The local population is equal to the sum of labor demands for all three employment types in all locations. This is due to the assumption that every individual in the second generation is in employment.<sup>7</sup> It is directly visible in Appendix A.2 that labor demand for both manual and abstract labor are higher in the more productive location, i.e. location 1. In order to determine location size, the crucial part is to identify where labor demand for routine occupations is higher. After a proof via contradiction, it is clear that the demand for routine occupations is higher in location 1 as well. Subsequently, as all labor demands are higher in the more productive location, this region is also larger in terms of population. Assuming the same area for all locations, a higher population size in a region translates directly in higher density in the same location.

**Proposition 2. Population Size** Assume  $\gamma_r < \theta$ .  $A_1 > A_2 \rightarrow S_1 > S_2$ , i.e. population in location 1 is larger than location 2. Given the assumption that area size is equal across regions, then location 1 exhibits higher density than location 2.

*Proof see Appendix A.4.*

---

<sup>7</sup>Due to the assumption of two members in each household (young and old) in every period, children do not change this model prediction.

Now I turn to the second dimension how intergenerational mobility is heterogeneous across space due to idiosyncratic demand for capital. From Proposition 1, I know that a rising capital stock in the economy reduces cross-generational transitions with respect to education, occupation and income. If the demand for capital is location-specific and related to the previous three properties, then the location demanding more capital will experience stronger polarization in educational choice than the location demanding less ICT capital.

**Proposition 3.** (*Polarization across Space*) Assume  $\gamma_r < \theta$ .  $A_1 > A_2 \rightarrow K_{1t} > K_{2t}$ , i.e. location 1 demands more capital, and therefore experiences stronger polarization in educational attainment than location 2.

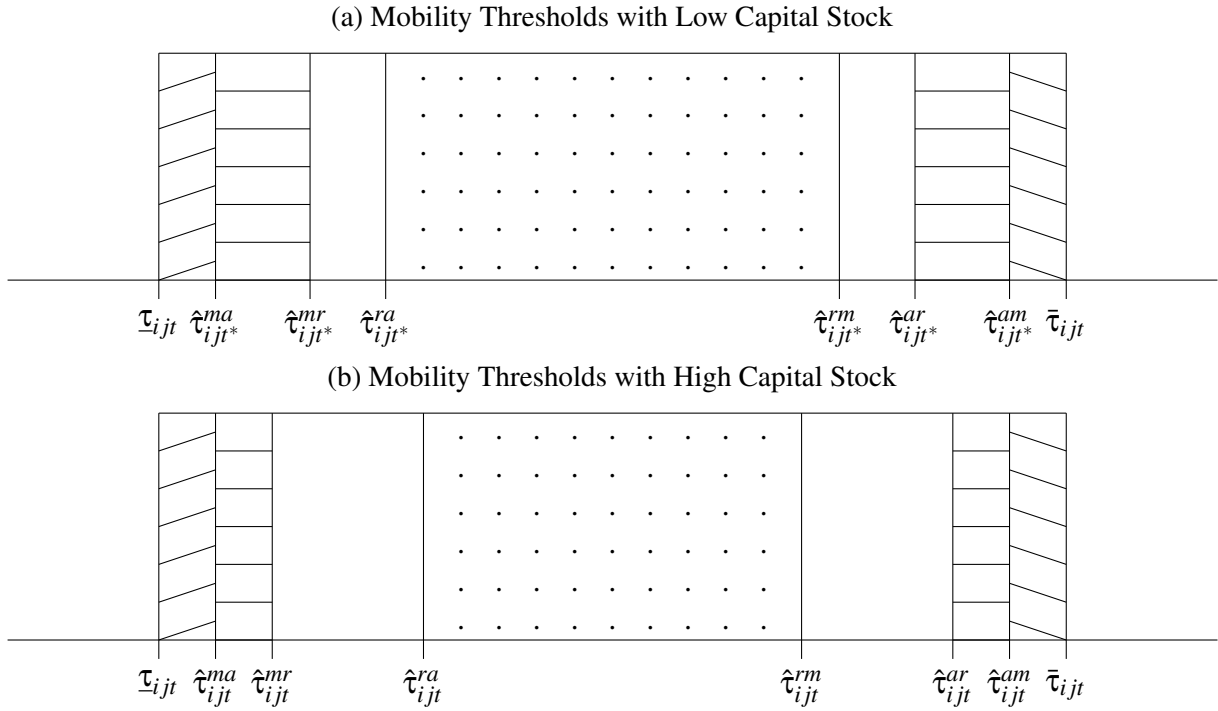
*Proof see Appendix A.5.*

The model predicts higher demand for capital for the more productive and more populous region. This is driven by the combination of substitutability between routine labor and capital and a larger absolute number of routine workers in the more populous location. The latter result formulated in Proposition (2) implies that there is greater potential for substitution between routine labor and capital in the larger location. In line with Davis et al. (2020), relative exposure is not the key driver that explains the destruction of middle-paid jobs in a given location. Instead, the falling price of computer capital is a necessary condition for the destruction of routine employment, but it is not sufficient. The sufficient condition is that there needs to be an incentive to destroy these jobs. In the model, the incentive to destroy middle-paid jobs depend on regional characteristics, primarily higher absolute advantages across all task types. The higher absolute advantage also manifests itself in housing prices, location size and thick tails.

Figure 1 shows the uniform distribution of educational frictions and the six thresholds of educational choice for two different states of the economy, one with a low and one with a high capital stock. The former can reflect either the economy with a high price for capital or a low-productivity location, while the latter can be interpreted either as the economy with a low price for the intermediate good or as a location with high high productivity. It reveals how the thresholds differ qualitatively based on the model predictions discussed above. The distribution of educational frictions ranges from its minimum value  $\underline{\tau}_{ijt}$  to its maximum  $\bar{\tau}_{ijt}$

First, I describe how upward mobility changes with respect to education with labor market polarization for children from parents working in the manual sector. Wages in routine occupations are falling, while wages in manual employment are unchanged. This creates an incentive to *not* invest in secondary education when the rise of computer capital is stronger. Hence, the threshold  $\hat{\tau}_{ijt}^{mr}$  in the lower panel of Figure 1 is shifted to the left compared to the upper panel. This finding implies that the marginal individual getting secondary education has lower educational frictions when the capital stock is low compared to when the capital stock is high. Consequently, upward mobility between low-income and middle-income jobs is *lower* in the latter, i.e. in the lower panel, compared to the former, i.e. the upper panel. In both panels, the area on the left with horizontal lines between  $\hat{\tau}_{ijt}^{ma}$  and  $\hat{\tau}_{ijt}^{mr}$  is the fraction experiencing upward

Figure 1: Education Thresholds



*Notes:* The figure depicts the thresholds of educational choice for two different levels of the of the capital stock. All six thresholds depend on parental bequests and future wage ratios. The upper panel shows the thresholds when the capital stock is low, and the latter when the capital stock is high. The thresholds indicate that more cross-generational transitions occur in the former compared to the latter. The areas with upward and downward sloping lines indicate the share of children who transition from manual to abstract and vice versa, respectively. They are unchanged in the model as capital does not impact their wage ratio. The areas with straight slopes indicate the children who transition from either manual or abstract parents into routine employment. These areas are larger in the upper panel as routine wages are falling less in a state with less capital. The dotted area indicates the share of children who do not experience cross-generational transfers out of secondary education. For the same reason, this area is also larger in the upper panel.

mobility from manual to routine, while the area right of the threshold  $\hat{\tau}_{ijt}^{mr}$  shows the share of children taking up jobs in manual occupations because they did not afford more than basic education. Therefore, the share of children from parents in manual workers also work in manual occupations.

Children with parents in the manual occupations can also exhibit upward mobility if they attain tertiary education and earn the respective future wage. The threshold depends on how wages in manual and abstract occupations change relative to one another. However, as argued above, the wage ratio between manual and abstract employment does not change with respect to the capital stock. Hence, threshold  $\hat{\tau}_{ijt}^{ma}$  is the same in both panels. Specifically, the area between  $\tau_{ijt}$  and  $\hat{\tau}_{ijt}^{ma}$  with the upwardly sloped lines indicates the share of children experiencing upward mobility from manual to abstract occupations. If computer capital had an impact on wages for both types of employment, then it would depend on the relative impact. Typically, and in line with skill-biased technological change, abstract workers would profit relatively more than manual workers. Hence, the wage ratio between manual and abstract employment would fall, and



by implication shift the threshold  $\hat{\tau}_{ijt}^{ma}$  to the right. However, higher costs  $z$  for tertiary education can counteract this shift to the right.

Second, I examine the changing thresholds for children with parents working in routine-intensive occupations, i.e. from the “middle class”. If individual educational frictions  $\tau_{ijt}$  are lower than the threshold  $\hat{\tau}_{ijt}^{ra}$ , then a child will enter abstract employment, whereas if their individual cost is higher than critical value of  $\hat{\tau}_{ijt}^{rm}$ , then they enter manual employment. In the former case, they exhibit upward mobility, and in the latter they are downward mobile. With constant wages in manual and abstract employment and declining wages in routine occupations, both thresholds are shifted stronger in opposite directions when the capital stock is high. This means that children are *more mobile in both directions*. Specifically, children of parents in routine employment increasingly choose either primary or tertiary education if the capital stock is high. Both shifts are reinforced by falling bequests as equation (5b) shows that they are a function of routine wages for children with this background. Finally, the dotted area in both panels indicates the interval where children do not change the sector of employment relative to their parents, i.e. they stay in routine employment. This area is smaller in the lower panel where the capital stock is higher. The total amount of routine workers is the dotted area plus the two areas with horizontal lines. In the upper panel all areas are larger compared to the lower one, reflecting a decline in routine employment when the capital stock is high.

Third, the educational choice of children whose parents work in abstract occupations is analogous to children from parents in manual occupations, just vice versa. That is, they are more likely get tertiary education than secondary education for routine occupations in location  $j$  where the demand for computer capital is stronger than in location  $j'$ . This is primarily driven by falling routine wages, but also reinforced by rising parental investment. This shifts the threshold  $\hat{\tau}_{ijt}^{ar}$  further to the right in location  $j$ , indicating lower downward mobility for children whose parents work in abstract jobs. Analogously to the argumentation for threshold  $\hat{\tau}_{ijt}^{ma}$ , the threshold and  $\hat{\tau}_{ijt}^{am}$  is the same in both cities. However, if polarization had a direct positive impact on abstract wages, this threshold would be shifted to the right as well, reflecting lower downward occupational transitions from abstract to manual wages across generations.

## 2.4 Qualifying Predictions

The model also provides predictions beyond the three key propositions mentioned in the previous section. Specifically, they refer to housing prices and “thick tails”. For the former, equation (10) reveals that real wages, i.e. nominal wages net of housing costs, are constant across locations. Further, I know that nominal wages for all tasks depend positively on the degree of exogenous location-specific productivity from equations (5a) to 5c). Therefore, housing prices are higher in more productive locations. Appendix A.3 examines the exact relationship between the housing price ratio and the productivity ratio. The productivity ratio is larger than the housing price ratio to the power of its expenditure share. Due to the Cobb-Douglas structure of the

utility function for the second period of life shown in equation (8), the expenditure share for housing is equal to  $\beta$ .

**Proposition 4. *Housing Prices*** Assume  $\gamma_r < \theta$ .  $A_1 > A_2 \rightarrow p_{1t} > p_{2t}$ , i.e. housing prices in location 1 are higher than in location 2.

*Proof see Appendix A.3.*

Eeckhout et al. (2014) show that larger cities in the United States exhibit larger shares of extreme skills than smaller cities, which they phrase as incidence of “thick tails”.<sup>8</sup> Similarly, Autor (2019, Fig. 7, Panel A) shows that the share of manual workers has increased in more dense commuting-zones, especially between 2000 and 2015. Eeckhout et al. (2019) show that in a similar setting, the model leads to thick tails if  $\gamma_m = \gamma_r = \gamma_m = \gamma$ . In words, if the output elasticity of manual, routine and abstract labor are equal, then more productive and larger region 1 exhibits thick tails. With this assumption, the model treats both input factors analogously, i.e. they are symmetric in their labor demands across locations.

My theoretical framework predicts thick tails due to individual heterogeneity with respect to educational frictions, and subsequent sorting into different occupational groups. Going back to Figure 1, the share of workers in manual workers is set up the following: it is the sum of all children from parents working in manual employment and with educational frictions larger than  $\hat{\tau}_{ijt}^{mr}$ , all children from parents working in routine employment and with educational frictions larger than  $\hat{\tau}_{ijt}^{rm}$ , and all children whose parents work in abstract employment and with educational frictions larger than  $\hat{\tau}_{ijt}^{am}$ . Comparing the upper and the lower panel, the figure reveals that the first two addends are larger when the capital stock is higher. Therefore, the share of workers in manual occupations is rising with an increasing capital stock. The same holds for abstract employment.

**Proposition 5. *Thick Tails*** Assume  $\gamma_r < \theta$ . Educational sorting due to individual heterogeneity in educational frictions ( $\tau_{ijt}$ ) leads to thick tails when the capital stock is high. This is driven by declining returns to secondary education and subsequent sorting into “extreme” educational attainment levels.

Economically, the incidence of thick tails means that abstract workers profit substantially from the presence of manual workers for two reasons. First, abstract workers use administrative work or other services provided by manual workers. Second, abstract workers demand manual workers via consumption spillovers, see e.g. Manning (2004). In this framework, low- and high-skilled workers do not complement in each other in production of a single good, instead low-skilled workers do “housework” activities for high-skilled workers, but these activities require physical proximity. Independent of the underlying explanation, the model predicts colocation of “extreme” skills in large cities.

---

<sup>8</sup>Davis et al. (2020) show that polarization does not occur symmetrically in France. Specifically, they show that the share of manual workers increases more in small cities, whereas the share of abstract labor is stronger in large cities.

### 3 Data Sources

This section presents the various data sources used in the subsequent empirical analysis which allow me to confront the model predictions with the data. I will first present the data I exploit for the commuting-zone analyses, for which the geographic dimension of the model takes on a major role. Then I turn to the Panel Study of Income Dynamics, which I exploit for variation over time and allows me to control for parental background.

#### 3.1 Commuting Zone Data

I exploit four different data sources in order to investigate whether the data confirms the model predictions. The main data source for multiple tests, such as ICT investment, educational polarization and upward mobility, are the Decennial Censuses available from IPUMS. Specifically, I use the censuses from 1970 to 2000 and the American Community Surveys (ACS) from 2010 and 2018 (Ruggles et al., 2018).<sup>9</sup> The data includes between 1% and 5% of the whole American population on all labor market statistics and more. From this data, I collect information on labor market polarization and other labor market variables, e.g. routine employment shares, similar to Autor and Dorn (2013), and on educational polarization. The authors highlight that there is no established measure of labor market polarization, and exploit the share of non-college service employment. I focus on the decline in routine employment, but also show results for upward mobility on non-college service employment. I compute changes in employment structure between 1990 and 2010 as the income in Chetty et al. (2014) is measured in the two years after 2010, and use the change in non-college service employment between 1990 and 2010. The data for the covariates also largely stems from the same data sources.

For upward mobility, I exploit the data on intergenerational mobility by Chetty et al. (2014). The authors estimate the relationship between parental income rank and children's income with a rank-rank specification for children born in the years 1980 and 1982. For their preferred measure, absolute upward mobility, which I also use in the benchmark regressions, they just need two parameters of their estimation, namely the slope and the intercept. They focus on children whose parents are located at the 25<sup>th</sup> percentile of the national income distribution and estimate the expected rank of these children. To estimate intergenerational mobility they use administrative records of more than 40 million children and their parents, and provide the data for all commuting zones with more than 250 children. One strong advantage of the data is that there is no variation in method nor collection dates, income is collected in 2011 and 2012, i.e. when the children are about the age of 30 to 32. They show that intergenerational mobility exhibits significant variation across regions, ranging from levels of the most mobile countries to below that of any developed country. Two further measures the authors provide are relative mobility, which measures the difference in outcomes between children from top versus bottom income families, and the transition probability where parents are in the lowest quintile of the national

---

<sup>9</sup>I cannot use the 2010 Census because it does not include key variables, e.g. on education.

income distribution and children end up in the highest quintile. Both measures are used in robustness analyses.

In order to measure density by commuting zone in 1970, I use the County Intercensal Tables 1970 to 1979 by the Census Bureau, which provide county-level population by various demographics, such as age, sex and race. In a given census years, the information on population comes directly from data collection. For the years in between two censuses, the Census Bureau estimates county population. This makes the use of 1970 more reliable than any other year from the County Intercensal Tables. I merge information on total population on the county level with land area in square miles in 1970 on the commuting-zone level. The next step comprises aggregating the county-level on population and land area to commuting zones defined by their 1990 commuting patterns with the help of the concordance table by the US Department of Agriculture. After the aggregation, I compute the (log) density for each commuting-zone. Density in 1970 varies substantially across commuting zones. The average density of a commuting-zone is just above 108 persons per square mile in 1970. In a commuting-zone at the 10<sup>th</sup> percentile, slightly above 13 people live within a square mile, while at the 90<sup>th</sup> slightly more than 930 live within the same area. This numbers are somewhat larger than in the overall population, but this is likely to be true because the smallest (and hence probably least dense areas) are not part of the sample as Chetty et al. (2014) require a minimum number of observations within a commuting zone.

Panel A in Table 1 shows descriptive statistics for the main variables crucial to the analysis and graphical representation. The first three measures give an overview of the variation of intergenerational mobility in the United States and are taken from Chetty et al. (2014). Their preferred measure, i.e. absolute upward mobility, measures the expected rank of a child whose parents are located at the 25<sup>th</sup> percentile of the national income distribution. The average expected rank is 42.57, so there is upward mobility on average, but as discussed already by Chetty et al. (2014) - considerable heterogeneity across space, seen by the standard deviation and both the 10<sup>th</sup> and 90<sup>th</sup> percentiles. For relative mobility, it is important to keep in mind that the higher the number, the lower upward mobility (see explanation in robustness analysis). In terms of extreme-quintile mobility, i.e. 8.7% of children whose parents are located in the bottom quintile of the national income distribution land in the highest quintile. This measures also exhibits strong heterogeneity, with the 90<sup>th</sup> percentile about 2.5 times as high as the 10<sup>th</sup> percentile.

The table also reports the degree of labor market polarization between 1990 and 2010. On average, the share of routine employment fell by 6.45% over this time period, with substantial heterogeneity ranging from more than 9.7% to a bit more than 2.2% shown by the 10<sup>th</sup> and 90<sup>th</sup> percentiles, respectively. The rise of each non-college and college service employment is similar on average, with 2.62% and 3.40%, respectively. These changes seem to indicate that a small share of the working-age population, i.e. about .5%, which was previously working in routine occupations, is not in employment because the sum of the means of the changes in the extreme skills are less than the loss in routine occupations. Further, there is strong regional vari-

ation in all three employment shares relative to working-age population. Similar to Autor and Dorn (2013), I exploit historical share of routine employment as an instrumental variable when I estimate the causal impact of labor market polarization on upward mobility. The table shows that the average share of routine employment in 1980 relative to total population is 33.63%.

I compute investment in information and communication technology (ICT) per worker on the commuting-zone level using and relate this to density and the share of population with a college degree, i.e. the “thick tails” prediction of the model. The Survey of Current Business is a quinquennial survey and it contains data on the distribution of new structures, equipment, and software from capital flow tables. However, the data does not include any geographic information, instead it contains detailed sectoral information, which I use to match it with sectoral information from the Census PUMS samples 1980 (for 1982) and 1990 (for 1992 and 1997). ICT investment per worker is measured in thousand USD with 2010 as base year. Between 1982 and 1997, average ICT investment per worker more than quadrupled in real terms. Average investment per worker rose from 8,140 USD in 1982 to 33,630 USD in 1997. Similar to the share of college degree, the commuting zone at the 10<sup>th</sup> percentile invests about half as much per worker as the commuting zone at the 90<sup>th</sup> percentile. This holds for all years, e.g. in 1997 the former invested 26,330 USD in ICT, while the latter spent 39,040 USD. The share of the population with a college degree in 1970 is equal to 6.08 percent, but also exhibits strong variation across commuting zones. The share of college degrees in the commuting zone at the 10<sup>th</sup> percentile with 3.82% is less than half of that in the commuting at the 90<sup>th</sup> percentile with 8.82%.

Panel B of Table 1 presents summary statistics for the panel data on the commuting-zone level used in the subsequent analysis. It shows the decadal decrease in routine employment and various measures of educational polarization between 1970 and 2018. The decrease in routine employment relative to working-age population by decade is larger than 3 percentage points, which is close to half the average value in Panel A, where the difference is 20 years. For educational polarization, I compute the five measures of educational polarization, namely the index by Allison and Foster (2004) and the indices with differential weights on the lower or upper part of the distribution by Naga and Yalcin (2008). As discussed in more detail below, one drawback applies to these measures. The values of the same index can only be compared for two different distributions if they have the same median. As Allison and Foster (2004) explain, the median is the central point of each of these measures, and serves as the reference point. Each measure takes into account the spread away from this median, and if the center changes, the spread will ultimately change, and the indices are not comparable any longer.

Due to the median-based approach of each education polarization index, the panel of commuting zones is unbalanced. This is because I have chosen to use the information on educational polarization from commuting-zones where the median education coincides with the national mode in a given year in order to get a maximum of observations with comparable polarization indices. The unbalanced panel is a result of varying medians of across commuting-zones and

Table 1: Commuting-Zone Summary Statistics

	(1)	(2)	(3)	(4)	(5)
	Obs.	Mean	Std. Dev.	P10	P90
Panel A - Cross-Sectional Data					
Absolute Upward Mobility	693	42.57	4.56	37.33	47.98
Relative Mobility	693	33.38	5.87	25.00	40.07
Transition Probability (P1, K5)	693	8.70	3.77	4.90	12.43
$\Delta$ Routine Emp.	693	-6.45	3.01	-9.72	-2.21
$\Delta$ Non-College Svc. Emp.	693	2.62	1.37	0.91	4.04
$\Delta$ College Svc. Emp.	693	3.40	2.29	0.54	6.32
Log Density (1970)	693	4.69	1.69	2.58	6.84
Share College Degree (1970)	693	6.08	1.86	3.82	8.28
ICT Investment per worker (1982)	693	7.06	1.54	5.11	8.90
ICT Investment per worker (1992)	693	8.14	2.37	5.48	11.13
ICT Investment per worker (1997)	693	16.60	4.90	11.08	23.17
Routine Employment (1970)	693	36.71	6.71	27.29	45.41
Panel B - Panel Data					
$\Delta$ Routine	2327	-3.02	2.93	-6.41	0.62
Allison-Foster	2327	2.85	0.47	2.22	3.48
ANY(2,1)	2327	0.30	0.06	0.23	0.38
ANY(1,2)	2327	0.35	0.05	0.27	0.41
ANY(4,1)	2327	0.39	0.09	0.28	0.51
ANY(1,4)	2327	0.43	0.06	0.35	0.50

*Notes:* The table shows summary statistics for commuting-zone variables used throughout the analysis. Panel A presents variables used in the cross-section. Cross-sectional data is used for presenting evidence of ICT investment and how it differs across commuting-zones, and for the final analysis relating labor market polarization and upward mobility. Panel B shows the decadal changes in routine employment and various metrics of education polarization between 1970 and 2018 using Decennial Censuses. The number of observations for the panel analysis is limited to those commuting zones where median education is equal to the national median as these metrics require the same median for comparability.

non-linear changes. Decadal changes in routine employment are measured from 1970 to 2018, i.e. the first difference is between 1980 and 1970, and the last between 2018 and 2010. Overall, in the United States there are 741 commuting zones, which would yield a total number of observations of 3705 if the panel was balanced. However, for the analysis I can only exploit 2327 commuting-zones, which means that I make use of around 465 commuting zones every decade in the panel.

Finally, the table shows the summary statistics for the five different measures of educational polarization. At a first glance, the previously-mentioned scaling between zero and one of the ANY( $\alpha$ ,  $\beta$ ) measures is apparent. Average polarization is lower if more weight is put on the upper part of the distribution reflected by a higher value of  $\alpha$ . This can be due to the overall

upskilling in the United States over the past decades as documented widely in the literature, e.g. by Castro and Coen-Pirani (2016). All indices also exhibit lower (absolute and relative) standard deviation in all educational polarization compared to the decline in routine employment relative to working-age population.

### 3.2 Panel Study of Income Dynamics

The Panel Study of Income Dynamics (PSID) is the longest-running panel household survey in the United States and allows me to follow individuals over their lifetime and retrieve information about their parents with respect to income. The survey started in 1968 and was conducted annually until 1997, and since then biennial, with 2017 being the latest available data. I focus on the ‘core’ sample of the PSID. I focus on children who later become heads (sons and daughters) or spouses (typically daughters). I use the family and individual family codes provided by the PSID to follow sons and daughters when they leave the parental household and form their own family unit.

As is standard in the literature of intergenerational mobility, I measure total family income, which includes taxable incomes and transfers received by the head, the spouse and other family members. It is consistently included in the PSID with nominal values.<sup>10</sup> To account for changes in nominal income due to inflation, I transform income into 2010 US dollar by using the average consumer price index (CPI) from the Federal Reserve Bank. I account for outliers by dropping the lowest and highest percentile of parental total incomes.

Solon (2002) argues that it is important to exploit the permanent income of the parents during teenagehood of children in order to estimate intergenerational elasticity. In other words, using only a single observation of parental total income during the age of 13 to 19 leads to the underestimation of intergenerational mobility because the parents’ income in a single year can be subject to transitory labor market shocks. To account for this, I average total parental income when children are between 13 and 19, provided there are at least three observations of parental income during teenagehood. Income of children can also be subject to transitory shocks, that is why I mean out the children’s income during adulthood over three waves.

The PSID has been used in many previous studies of intergenerational mobility. A non-exhaustive list includes Solon (1992), who was the first to exploit the PSID as a data source to estimate intergenerational elasticity (IGE). Further, Lee and Solon (2009) show that more observations from the PSID can be used by introducing quartic polynomials around a centered age, which account for the life-cycle bias. Palomino et al. (2018) uses the PSID to estimate IGE along the parental income distribution. Mazumder (2018) provides a review of the literature and how the PSID contributed to the now widely spread view that the United States is among the least socially mobile countries among developed countries. The main alternative data set for intergenerational mobility in the United States is the National Longitudinal Survey of Youth

---

<sup>10</sup>As the interviews are conducted throughout the year, the incomes refer to the previous year of the interview.

(NLSY), which has been used by e.g. Kourtellos et al. (2020). Palomino et al. (2018, Table 4) provide an overview of IGE studies including of data sources and point estimates. Typically, the point estimates using the PSID range between 0.34 and 0.51 without demographic controls, depending on sample and time period used. Usually, education reduces this estimate considerably.

### 3.2.1 Descriptive Statistics

Based on the sample selection described above, the data set of children whose family income I can observe between 1980 and 2016. In total, my sample consists of 85,865 observations encompassing both sons and daughters, and where I can observe parental income during teenagehood. The total number of sons and daughters I observe is equal to 8493, implying that I observe each child of the parent-child pair on average 10.11 times. Of all children, I can also observe the occupations of all parents (if present), however in both the subsequent descriptive statistics and the analysis I will concentrate on the occupation of the head of the household, assuming that this is the parent with higher earnings during teenagehood.

Table 2: Summary Statistics of Panel Study of Income Dynamics

	Unemployed	Manual	Routine	Abstract
Family Income	10.56 (0.93)	10.84 (0.81)	10.90 (0.76)	11.16 (0.79)
Parental Income	10.39 (0.65)	10.73 (0.54)	10.87 (0.42)	11.22 (0.50)
Share Female	55.89 (49.65)	50.60 (50.00)	51.58 (49.98)	50.00 (50.00)
Age	37.31 (9.29)	38.32 (9.77)	38.91 (9.97)	37.99 (9.66)
Education	12.66 (2.06)	12.97 (1.93)	13.42 (1.99)	14.45 (2.03)
Share White	60.55 (48.88)	79.84 (40.12)	81.22 (39.05)	91.96 (27.19)
Share Black	35.73 (47.92)	16.03 (36.69)	15.78 (36.46)	5.56 (22.91)
Share Other	3.72 (18.92)	4.13 (19.91)	2.99 (17.04)	2.48 (15.54)
Observations	11927	26729	17269	29940

*Notes:* This table shows summary statistics of family incomes of both parents and children, and demographic characteristics of the children by occupational background of the family's head during the child's teenagehood. The summary statistics are means and standard deviation in brackets. Occupational background is divided into four distinct categories, i.e. unemployed, manual, routine and abstract. The latter three categories are of interest for the analysis below.

Table 2 shows the variable of interest, i.e. family income (in logs), by occupational background in three categories. These categories are unemployed, manual, routine and abstract. In what follows, I will concentrate on the last three categories because they reflect the occupations



in the task framework by Autor et al. (2003) incorporated in the model above. Unsurprisingly, parental income rises with occupational category, i.e. average parental income of parents who work in a manual occupation is smaller than those in routine occupations, which again is smaller than parents working in abstract occupations. This order has not changed in terms of family incomes of the children, but the gaps are somewhat smaller, while the standard deviation is much larger. One reason for the higher standard deviations could be larger within-occupation inequality as documented e.g. by Kambourov and Manovskii (2009) for United States. Besides the rise of superstar firms, e.g. Autor et al. (2020), the theory in section 2 offers implicitly an alternative explanation for this pattern: Due to limited intergenerational mobility, individual ability varies stronger for new generations entering different occupations.

Panel (a) of Figure 2 shows the quintiles of point estimates of intergenerational mobility across states based on equation (20) explained below. I show the results already here in order to show a comparison of regional social mobility with Chetty et al. (2014)[Fig. VI]. Two remarks are necessary for the comparison of their map and the map based on the PSID. First, public data in the PSID only provides geographic information on the state-level and not on a lower level, whereas Chetty et al. (2014) have administrative data, i.e. tax returns, which allows the authors to provide more granular estimates of intergenerational mobility. Given the comparatively small number of observations in the PSID compared to the administrative data, the likely level of aggregation of the PSID would have been on the state level as well even with the availability of more granular geographic information.

Second, the map in Chetty et al. (2014) shows absolute upward mobility, which differs from intergenerational elasticity, which I am estimating in line with previous work. Chetty et al. (2014) argue that IGE combines the dependence features captured by the rank-rank slope with the ratio of income inequality measured by their standard deviation across generations. In order to account for this bias due to diverging income inequality across generations across states, I compute both the average standard deviation during teenagehood, i.e. the same years when I measure parental income, in the state where the child spent most years during the time and the standard deviation of the state where the child is living when I observe it during adulthood. With the point estimates for state-level IGE and measurements of income inequality across generations expressed by standard deviations, I can apply Chetty et al. (2014, Eq. 1):

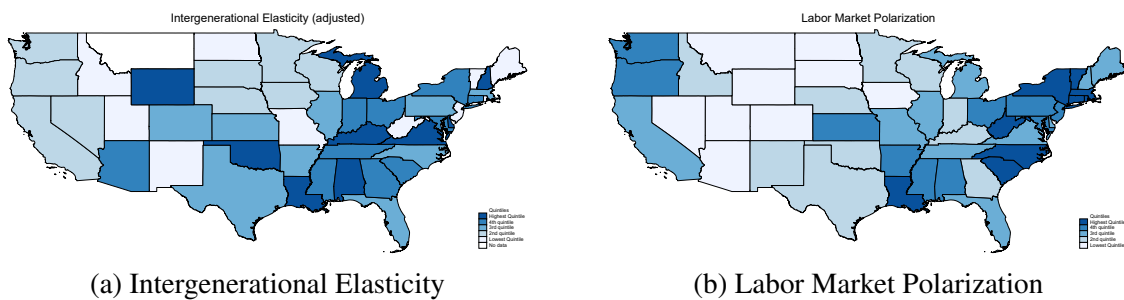
$$IGE = \rho_{XY} \frac{SD(\log Y_i)}{SD(\log X_i)}, \quad (17)$$

where  $\rho$  is the correlation between log child income ( $X$ ) and log parent income ( $Y$ ) and  $SD()$  represents the standard deviation of the respective generation, I obtain “adjusted” estimates of intergenerational mobility for the majority of states using the PSID.

The left panel of Figure 2 shows the regional distribution of adjusted intergenerational mobility in the United States in quintiles. Darker colors indicate that intergenerational elasticity is larger than in states with lighter colors. Similar to Chetty et al. (2014), I find that there is a higher

elasticity in the South and lower persistence across generations in the states at the West Coast and in the Midwest. One notable exception is Wyoming, which could be due to a low number of observations in this state (84 in total). The only state with not enough observations is Montana (shown in white). The map shows that even with a much smaller data set, the regional distribution of social mobility is similar to the map constructed with administrative data, though it lacks geographic granularity. The estimates used for the construction of the map do not account for demographic characteristics in order to establish comparability with the maps of Chetty et al. (2014) as the authors do not have these information, hence their maps only rely on rank-rank correlations.

Figure 2: Intergenerational Mobility by State



*Notes:* The left panel shows map shows the spatial distribution of the point estimates of intergenerational elasticity for all states adjusted for changes in inequality over time. Darker colors indicate higher point estimates, i.e. states with higher immobility. The spatial distribution is similar to that of the commuting-zones in Chetty et al. (2014)[Fig. VI]. The Southeast is on average much more immobile than other regions in the United States. The right panel shows the rise of non-college service employment between 1990 and 2010 as a proxy for labor market polarization.

Panel (b) in the same figure shows the degree of labor market polarization proxied by the rise in non-college service employment between 1990 and 2010. The two maps show strong spatial correlation between labor market polarization and intergenerational elasticity. As argued above for the model, intergenerational elasticity for the two “extreme” skills, i.e. manual and abstract, should go up with labor market polarization, whereas IGE should decrease for children whose parents work in routine occupations. The overlap of the maps seems to indicate that the tails of the occupational groups dominate the middle.

## 4 Empirical Evidence

In this section I confront the key model predictions with the data. First, I provide empirical evidence that labor market polarization reduces intergenerational mobility across space and time. For the former, I exploit data on absolute upward mobility from Chetty et al. (2014), which computes the expected rank of children in the national income distribution from parents who are located at the 25<sup>th</sup> percentile of the national income distribution. For the development across time, I make use of the PSID data and determine whether intergenerational elasticity

changes over time for children where the heads are working in manual, routine or abstract occupations.

The main channel proposed in the model is educational choice. The model predicts that educational attainment of young labor market entrants becomes more polarized as wages for routine occupations decline. I start by showing that education became more polarized in the United States between 1970 and 2018 using Census data among the group of young people between 20 and 29 and in the labor force. I also compute “family premia” for various educational levels depending on the educational level of parents. The next step encompasses to estimate the impact of changes of the decline in routine employment educational polarization. The last part of education exploits the Educational Supplement from the Current Population Survey, and allows me to investigate how educational choice depends on the interaction of past labor market polarization and family background.

In order to lend further support to the model, I conclude the empirical analysis with evidence that the “qualifying predictions” from the model in Section 2 hold in the data. Specifically, I focus on density, ICT investment, and housing prices.

#### 4.1 Upward Mobility

In this section I estimate the causal impact of labor market polarization on intergenerational mobility exploiting geographic variation in the expected rank of children from low-income families. Chetty et al. (2014) provide the data for the birth cohorts from 1980 to 1982. This means that they enter the labor market around the year 2000. The authors measure the income rank of these birth cohorts around the age of 30. I proxy labor market polarization as the decline in routine employment relative to working-age population on the commuting-zone level. This implies that a *higher value* of the proxy implies *weaker* labor market polarization.

##### 4.1.1 Empirical Specification

Endogeneity concerns can arise for two reasons. First, omitted variables bias can distort the point estimates. A region’s capacity or willingness to absorb new ideas and technology independent of the background of the inventors might be correlated with both labor market polarization and intergenerational mobility. A higher degree of absorption may increase polarization because new ideas typically spread first within the same region. Aghion et al. (2018) shows that innovation by new entrants relates positively with upward mobility. The positive correlation between both the dependent and independent variable introduces a downward bias.<sup>11</sup> The second bias can arise from measurement error. It is crucial to remember that data on upward mobility provided by Chetty et al. (2014) are point estimates with a standard error, i.e. the

---

<sup>11</sup>Normally, a positive correlation of the omitted variable with both the dependent and independent variable causes an upward bias. But as I proxy labor market polarization with a change in routine employment, and a higher value implies lower polarization, a downward bias is the consequence.

measures are not actually observed, but inferred from administrative records. This introduces some potential measurement error, implying a bias towards zero if the error is random.

To encounter the endogeneity concerns, I estimate the causal impact of labor market polarization on upward mobility using an instrumental variable (IV) estimation strategy. The instrumental variables I use to estimate the causal impact of labor market polarization on the intergenerational mobility builds on the model in Section 2. Specifically, I use both (log) density of a commuting-zone and the share of routine-intensive employment in 1970.<sup>12</sup> I exploit variation in 1970 as the arrival of the first three commercially successful computers occurred in 1977, the so-called “1977 Trinity”. Based on the model, an alternative to log density as instrumental variable is total population. However, density seems a more appropriate measure compared to population as interactiveness and proximity for both manual and abstract tasks have been highlighted in the literature.

In order to identify the causal impact with an IV estimation, two conditions have to be met. First, the instrument must induce a change in the endogenous variable as theory predicts. Based on previous findings and the model outlined above, this means that historical routine task-intensive employment and log density must have a negative impact on the change in routine employment in the first stage. The second condition of a valid IV is the exclusion restriction, in particular the instruments are not allowed to have an effect on upward mobility other than through its effect on polarization conditional on covariates.

The identification strategy to estimate the causal impact of labor market polarization on upward mobility takes on the following form:

$$IGM_{js} = \beta_0 + \beta_1 \Delta Routine_{js} + \mathbf{X}'_{j2000} \Theta + \gamma_s + \varepsilon_{js}, \quad (18)$$

$$\Delta Routine_{js} = \alpha_0 + \alpha_1 \log(Density)_{j1970} + \alpha_2 RoutineShare_{j1970} + \mathbf{X}'_{j2000} \Theta + \gamma_s + e_{cs} \quad (19)$$

where  $IGM_{js}$  denotes the expected rank of children in the national income distribution from parents from the 25<sup>th</sup> percentile in commuting zone  $j$  in state  $s$ .  $\Delta Routine_{js}$  is the change in routine employment between 1990 and 2010 because the income rank of the baseline cohorts from Chetty et al. (2014) is measured between 2011 and 2012. The vector of controls  $\mathbf{X}'_{j2000}$  contains covariates from the main specification in Autor and Dorn (2013, Table 5) from the year 2000 in order to account for labor market conditions when the baseline cohorts in the data on upward mobility by Chetty et al. (2014) enter the labor market. It includes various demand and supply shifters, specifically the ratio of college to non-college educated individuals in the population, the unemployment rate, the share of employment in manufacturing, the elderly share of population, the female labor force participation rate and the start-of-decade fraction of non-college workers in a commuting zone whose real wage is below the minimum wage

---

<sup>12</sup>I show below that historical log density is a good predictor for ICT investment per worker and that the relationship is increasing over time.

that will be enacted in the subsequent decade.<sup>13</sup> The regression also includes state-dummies ( $\gamma_s$ ). Observations are weighted by the commuting zone share of national population in 1980. Standard errors are robust to heteroscedasticity.

The parameter of interest is  $\beta$  in equation (18), which exhibits a *positive* sign if labor market polarization has a detrimental impact on absolute upward mobility. Potentially, the impact could be positive due to upward mobility if children from low-income parents leap over the middle of the income distribution and enter abstract high-paying occupations. In the model, this does not occur at a higher rate with ongoing labor market polarization because I assume symmetry for manual and abstract occupations. However, as Autor and Dorn (2013) show, the wage increase for high-skilled workers is larger than for low-skilled workers. Equation (19) shows that labor market polarization is regressed on the instrument, namely log density and the share of routine employment in 1970. The important parameters in the first stage represented in the equation are  $\alpha_1$  and  $\alpha_2$ . Based on the model and findings by both Davis et al. (2020) and Eeckhout et al. (2014), the first parameter should be negative for the change in routine occupations. The sign should be equivalent for the historical share in routine-intensive occupations based on Autor and Dorn (2013).

#### 4.1.2 Results

Table 3 provides evidence that labor market polarization reduces absolute upward mobility. The first two columns present the results for the ordinary least square (OLS) estimator. The first column does not include any controls nor state-fixed effects, whereas the latter one includes both. Column (3) displays the causal effect of labor market polarization on upward mobility based on the IV identification strategy with historical log density and historical routine employment as instrumental variables. The lower panel shows the first stage results. Overall, the results suggest a negative relationship between intergenerational mobility and labor market polarization. A less strong decline in routine employment, i.e. weak labor market polarization, increases the expected ranks of children of parents at the 25<sup>th</sup> percentile of the income distribution. This relationship is both economically and statistically meaningful. As argued above, the point estimates are likely biased towards zero, which is reflected in the difference between the OLS and IV estimates.

The point estimate in column (3) of the first panel indicates that an increase of one percentage point in labor market polarization leads to a decline of .57 expected ranks of a child whose parents are located at the 25<sup>th</sup> percentile of the income distribution. This result is economically meaningful as the median commuting zone lost 7.70 percent of routine jobs relative to the working-age population. Extrapolating out of sample therefore implies that a child moving from the median commuting-zone in terms of routine employment decline to a commuting zone without labor market polarization is expected to be located 4.4 ranks higher in the na-

---

<sup>13</sup>I exclude the share of the non-college population that is foreign born because upward mobility is substantially higher for migrants (see e.g. Abramitzky et al. (2019)).

tional income distribution. According to Chetty and Hendren (2018b), a one percentile increase corresponds to a higher income of approximately 818 USD. Taking the average income of children with parents earning below median from the same authors, i.e. 26,091 USD, a move as described before translates into 13.79% higher income.

Table 3: Effect of labor market polarization on upward mobility

<b>I. 2<sup>nd</sup> Stage</b>	(1)	(2)	(3)
	OLS	OLS	IV
$\Delta$ Routine Emp.	0.77*** (10.63)	0.23*** (4.41)	0.57*** (5.38)
State FE	No	Yes	Yes
Controls	No	Yes	Yes
Obs.	693	693	693
R <sup>2</sup>	0.26	0.75	
F-Statistic			122.32
<b>II. 1<sup>st</sup> Stage</b>			
Log Density (1970)			-0.40*** (-4.36)
Routine Emp. (1970)			-0.24*** (-12.29)
Partial R <sup>2</sup>			

*Notes:* The dependent variable of the second stage of the IV approach is absolute upward mobility regressed on labor market polarization measured as the change in routine employment between 1990 and 2010. The lower panel shows the first-stage results. Control variables and state fixed effects are not shown. Observations are weighted by population share in 1980. t-statistics are shown in brackets. \* denotes 10% significance, \*\* denotes 5% significance, \*\*\* denotes 1% significance.

The second panel of the same table shows the estimation results of the first-stage, i.e. equation (19). Based on the model, (log) density should impact the decline in routine occupations and the rise in manual workers negatively and positively, respectively. The underlying idea is that the more dense region is demanding more capital, hence leading to stronger automation of routine tasks and, consequently, experience stronger labor market polarization. The results suggest that the prediction from the model holds true, an increase of population density by 1 percent is associated with an decrease of routine employment by .40 percentage points. Unsurprisingly, historical specialization in routine employment exhibits the same sign. In line with expectations and the authors findings, this point estimate is also positive and at a similar scale in the OLS estimations in Autor and Dorn (2013, Table 5, Panel A). All point estimates are statistically significant at all conventional levels and the F-Statistics is far above the value of 10. Table 4 extends the benchmark analysis by introducing further control variables, which are associated with upward mobility. Based on previous research, Chetty et al. (2014) sort them into nine different categories plus the fraction of black residents. Because all nine categories encompass three covariates, I exploit the first principal component of each category. The original data is taken from Chetty et al. (2014) and usually measures the variables around the year 2000,

i.e. when the benchmark cohorts enter the labor market. This coincides with the timing of labor market controls included in vector  $\mathbf{X}'_{j2000}$  in the baseline regressions.

First, migration (MIG) encompasses the fraction of foreign born and both migration inflows and outflows. Second, labor market conditions (LAB) contain the share of manufacturing, Chinese import growth and the teenage labor force participation rate. Third, college education (COL) measures colleges per capita, college tuition and college graduation rate. Fourth, local tax policies (TAX) includes local tax rates, state EITC exposure and tax progressivity. Fifth, family structure (FAM) involves the fraction of single mothers, the divorce rate and fraction married. Sixth, social capital (SOC) comprises a social capital index, fraction religious and the violent crime rate. Seventh, high-school education (K12) consists of student-teacher ratio, test scores and drop-out rate. Eighth, income distribution (INC) includes mean household income, the Gini coefficient and top 1% income share. Ninth, segregation (SEG) contains an index of racial segregation, segregation of poverty and the fraction commuting less than 15 minutes to work. Lastly, I add the fraction of black residents (BLA). Most variables in the first four categories exhibit weak correlations with absolute upward mobility, while the correlations are stronger for the variables in the last five categories.

Table 4 shows that the overall negative relationship between labor market polarization and upward mobility holds. As before, the sign of the coefficients is positive as I use the change in routine employment between 1990 and 2010 as a measure for labor market polarization. In most estimations, the coefficient for labor market polarization is largely unchanged compared to the point estimate of the IV point estimate in Table 3. Disregarding the specification with segregation (by income and race) in column (9) the point estimates range between .24 and .62. In all these specifications, the coefficient of labor market polarization is still statistically significant at the 1% level. The addition of the various covariates without a strong decline in the point estimate of labor market polarization indicates that labor markets play an important role for intergenerational mobility as suggested by Rothstein (2019).

Columns (3) and (7) in Table 4 include the first principal components of the covariates describing the education system, i.e. college and high-school education, respectively. Especially the specification including the principal component for high-school education exhibits a substantial decrease. This is indicative of the importance of the education system on upward mobility. The point estimates in columns (5) and (10) are also substantially smaller than the benchmark point estimate. They include family characteristics and the share of blacks, respectively. While the former exhibits strong negative correlations with upward mobility, this is likely driven by design. This is because Chetty et al. (2014) measure income on the family level, and therefore the share of single mothers is strongly negatively related to upward mobility. The result in column (5) is entirely driven by the share of single mothers.<sup>14</sup> However, Chetty and Hendren (2018b) show that the causal impact of family structure is small. With respect to the share of blacks,

---

<sup>14</sup>The point estimate for  $\Delta Routine$  when excluding the share of single mothers when constructing the principal component for FAM is .60, and not significantly different from the benchmark point estimate in Table 3.

Table 4: Effect of labor market polarization on upward mobility - Additional Controls

<b>I. 2<sup>nd</sup> Stage</b>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	MIG	LAB	COL	TAX	FAM	SOC	K12	INC	SEG	BLA
$\Delta$ Routine Emp.	0.92*** (6.11)	0.81*** (5.74)	0.75*** (5.77)	0.82*** (6.12)	0.41*** (3.77)	0.78*** (6.18)	0.58*** (3.64)	0.72*** (5.00)	0.06 (0.52)	0.67*** (5.40)
Obs.	691	693	563	693	693	666	663	693	693	693
F-Statistic	69.88	84.37	77.46	83.18	55.33	67.08	69.30	60.59	36.15	72.65
<b>II. 1<sup>st</sup> Stage</b>										
Log Density (1970)	-0.36*** (-3.01)	-0.41*** (-3.41)	-0.39*** (-2.90)	-0.39*** (-3.24)	-0.34** (-2.58)	-0.34*** (-2.70)	-0.23* (-1.82)	-0.03 (-0.22)	-0.26 (-1.56)	-0.30** (-2.22)
Routine Emp. (1970)	-0.25*** (-9.09)	-0.24*** (-9.73)	-0.24*** (-9.17)	-0.24*** (-9.32)	-0.24*** (-9.61)	-0.24*** (-9.36)	-0.23*** (-8.74)	-0.23*** (-9.75)	-0.23*** (-8.90)	-0.25*** (-9.75)
$R^2$	0.72	0.72	0.72	0.72	0.72	0.72	0.74	0.73	0.72	0.72

*Notes:* The dependent variable of the second stage of the 2SLS approach is absolute upward mobility regressed on labor market polarization measured as share of non-college service employment. Columns (1) to (9) include the first principal component of all variables included in Chetty et al. (2014, Figure 8). MIG stands for migration, LAB for labor markets, COL for college, TAX for the tax system, FAM for family, SOC for social capital K12 for high-school, INC for properties of the income distribution, SEG for segregation, and BLA the fraction of black residents. Control variables and state fixed effects are not shown. Observations are weighted by population share in 1980. t-statistics are shown in brackets. \* denotes 10% significance, \*\* denotes 5% significance, \*\*\* denotes 1% significance.



this result is potentially driven by lower levels of upward mobility for blacks compared to other races as shown by Chetty et al. (2020).

The main exception is SEG in column (9), where the point estimate of labor market polarization is not statistically significant and actually changes signs. The results does not seem to be driven by a specific variable included in the first principal component. The F-Statistic is also the lowest for this specification, though it is still comfortably above the critical value of 10, one explanation could be that segregation of poverty and/or race is stronger in dense areas. Boustan (2013) argues that educational attainment and earnings are lower for blacks in more segregated cities, while Quillian (2014) generalizes these findings for advantaged versus disadvantaged groups. Consequently, segregation potentially plays a crucial role for upward mobility as argued by Fogli and Guerrieri (2019).

In most specifications, the first-stage results are similar to the benchmark results, i.e. the coefficient for log density is typically larger than the point estimate for historical routine employment. Exceptions are INC and SEG, where the point estimate for log density is statistically insignificant, and close to zero for the former. This is unsurprising, however, as for example in more dense regions average income is higher, as is also predicted by the model in Section 2. Somewhat more surprising is that segregation seems to be an issue in more dense commuting-zones. The coefficient for specialization in routine employment stays also similar to its value in the benchmark specifications, but stays statistically significant in all specifications.

## 4.2 Intergenerational Elasticity

After determining whether the empirical evidence supports the key prediction that labor market polarization reduces upward mobility across space, I now turn to variation across time. Due to data constraints on absolute upward mobility over time, I exploit a different measure of intergenerational mobility, namely intergenerational elasticity (IGE). It is defined how strong parental income affects the children's income during adulthood.

The model predicts divergent patterns of intergenerational elasticity depending on parental background. Specifically, with progressing labor market polarization IGE should be higher for children with parents in manual and abstract occupations. This is driven by both less transitions from each of these groups into routine occupations and higher persistence to choose the same educational attainment as their parents. On the other hand, IGE should fall for children from parents in routine jobs as the rate of transitions increases into both directions, rendering parental income less important than for the other two groups. This measure allows me to investigate trends over time, and to investigate patterns of intergenerational mobility across various parental backgrounds. Precisely, I define the occupation of the household head during teenagehood as the parental background.

### 4.2.1 Empirical Strategy

Using the PSID data described in Section 3, I estimate the intergenerational elasticity for each group dependent on parental background. The drawback of using the occupation of the household head when the individual was a teenager is that I cannot measure whether the parent actually lost their job due to labor market polarization. Instead, the occupational status of the head serves as indicator of exposure to rising automation of routine tasks. In the baseline estimation, I differentiate between manual, routine and abstract occupations.

I follow the standard literature in terms of estimation strategy by accounting for the life-cycle bias of both parents and children. Adopting this strategy from Lee and Solon (2009) allows me to exploit the entire pool of available data, i.e. pairwise children and parents. The estimation equation takes on the following form:

$$\ln y_{c_{it}} = \beta_0 + \beta_1 \ln y_{p_i} + \beta_2 [\ln y_{p_i}] OccHead_i + X_i' \Omega + \sum_{n=1}^4 \delta_n A_i^n + \sum_{n=1}^4 \gamma_n C_{it}^n + \sum_{n=1}^4 \theta_n [\ln y_{p_i}] C_{it}^n + \varepsilon_{it}, \quad (20)$$

where  $y_{c_{it}}$  measures the real family income (in logs) of adult children from family  $i$  in year  $t$ . Similarly,  $y_{p_i}$  represents average real family income (in logs) when the child was between 13 and 19. The vector  $X'$  includes demographic variables such as education, race, marital status and a binary variable indicating whether the individual is a head or a spouse, and the occupation of the head during teenagehood. To account for the life-cycle bias of the parents,  $A_i$ , and parameters  $\delta_1$  to  $\delta_4$ , depict the age of the head in family  $i$  when the child was 16 years old. Equally important is to account for the life-cycle bias of children, I control for the fourth polynomial of child's age centered around the age of 40, represented by  $C_{it}$  and parameters  $\gamma_1$  to  $\gamma_4$ . This is in line with Black and Devereux (2011) who show that income around the middle of the life-cycle is the best proxy of permanent income. The polynomial of child's age (centered around 40) is also interacted with parental income during teenagehood in order to account for potential divergence in income patterns across the life-cycle depending on parental income.

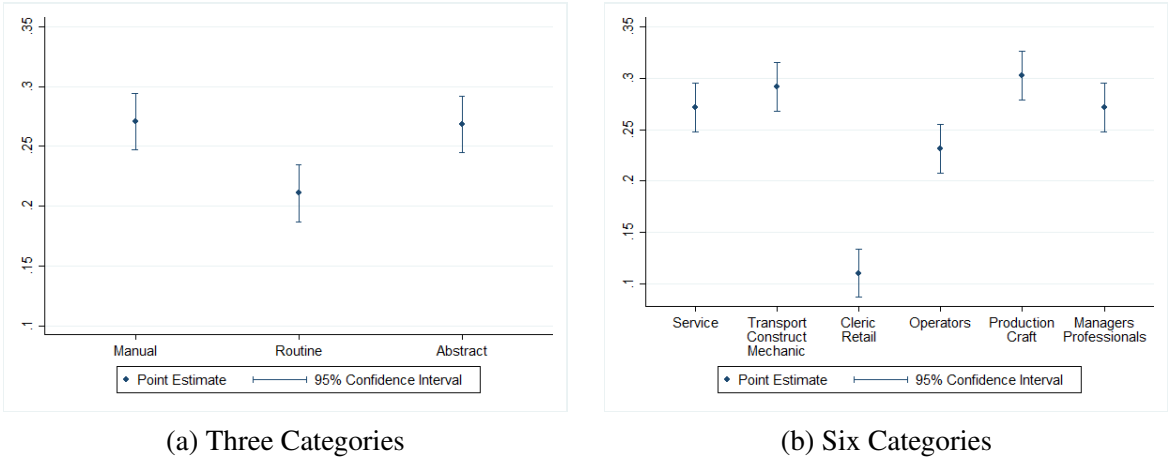
The main interest lies in the linear combinations of the estimated coefficients  $\beta_1$  and the vector  $\beta_2$ . As the modifying variable  $OccHead_i$  is a categorical variable, I can simply sum up the coefficients of interest as laid out in Brambor et al. (2006), and in order to compute meaningful standard errors I use the equations by Aiken et al. (1991). To get consistent estimates across time, I harmonize occupation data in the PSID with the help of IPUMS crosswalks as the PSID uses of occupational codes from the Decennial Censuses.

### 4.2.2 Results

Figure 3 presents how intergenerational elasticity differs depending on occupational background of the head. The left panel differentiates occupations by three groups, namely manual, routine and abstract. This is equivalent to the task framework by Autor et al. (2003), and the right panel

broadens the number of occupational groups to six, following Autor and Dorn (2013, Table 2). I have sorted the six groups along the more coarse three occupational groups according to the findings of the authors. The first two of the six groups are manual occupations, while the next three groups are more routine-intensive. Finally, the last subgroup is analogous to the group of abstract occupations.

Figure 3: IGE by Occupational Background



Notes: The left panel shows the point estimates of intergenerational elasticity (IGE) based on the three occupational categories referred to in the task framework in Autor et al. (2003), i.e. manual, routine and abstract. A u-shaped pattern is visible, indicating that parental income is more important for the child’s income if they are coming from the lower or upper part of the occupational structure. The right panel breaks up these three broad categories into six categories. The first two categories belong to manual occupations, the following three belong to routine occupations, and the last coincides with abstract occupations.

Both panels of Figure 3 exhibit a u-shaped pattern of intergenerational elasticity depending on occupational background. This provides evidence that parents influence their children’s income stronger if they work in a manual or an abstract job, compared to parents who work in routine employment. IGE is essentially equal for children with parents working in manual and abstract occupations with a value slightly above .27. IGE is lower for children with parents working in routine-intensive occupations with a value of  $\approx .22$ . The 95% confidence intervals indicate that the point estimates for manual and abstract are statistically different from routine. The point estimates are quite low compared to previous studies because I include a wide range of demographic characteristics as covariates.

The right panel with the slightly more granular division of occupational groups gives some more insights, specifically whether any particular subgroup drives the results. Interestingly, the intergenerational elasticities if parents work in any subgroup within manual occupations, i.e. service or transport/construction/mechanic, are statistically not different from one another. The low level of IGE within the group of routine occupations is particularly driven by the occupation group of clerics and retailers, a group strongly associated with routinization and labor market polarization. The point estimate is below .15. On the other hand, children whose parents work

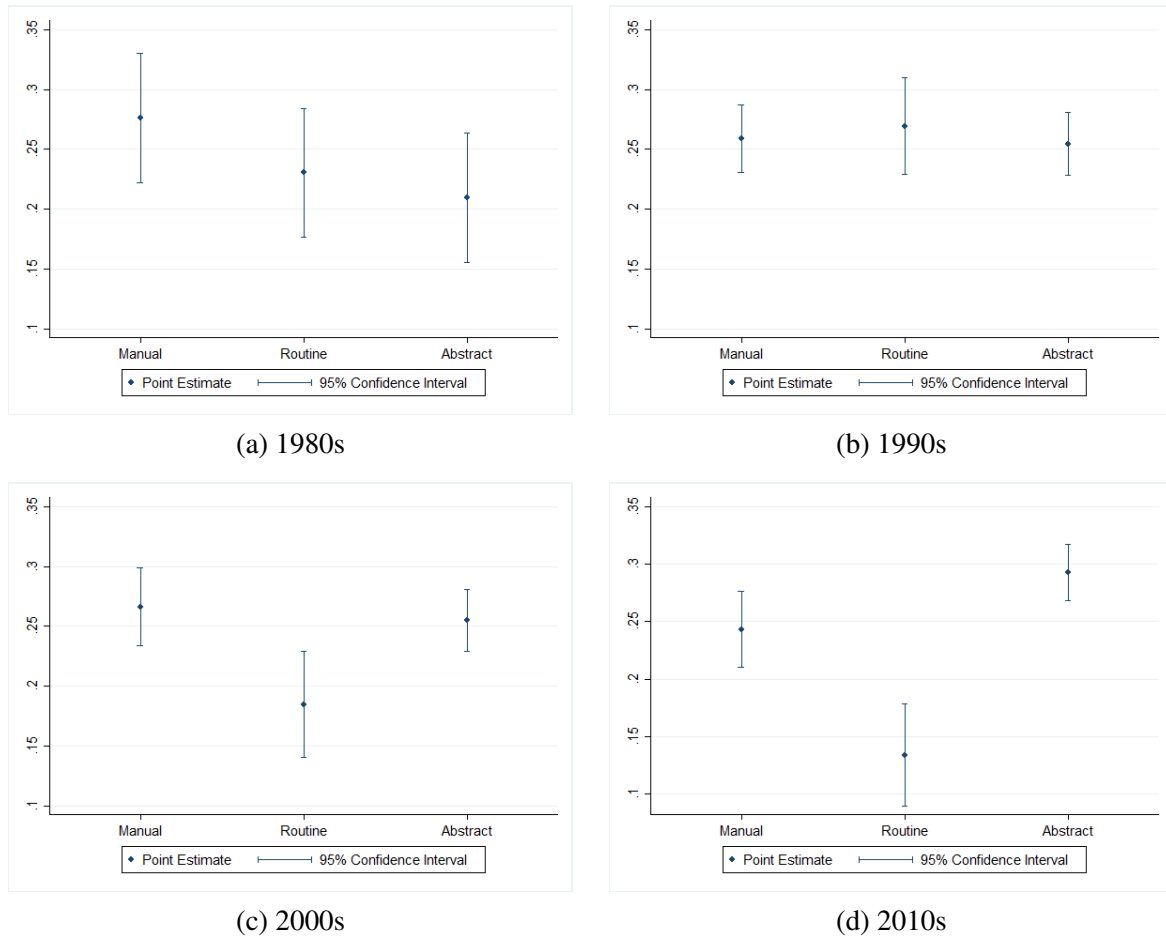
in production or craft, which is the highest paying routine occupation, the IGE slightly surpasses that of abstract occupations, albeit they are not statistically different from one another. Obviously, the point estimate and the standard errors for the final more granular subgroup are equivalent to that of abstract ( $\approx .27$ ) in the left panel as the group of abstract occupations is not divided into multiple groups.

Besides the investigation of various family backgrounds, the PSID also allows me to extend the baseline estimation to examine whether the levels of intergenerational elasticity have changed over time. I estimate equation (20) where I extend the modifying variable, i.e. *OccHead*, to the interaction between occupational group and decade, and include decade-fixed effects. The triple interaction consisting of parental income during teenagehood, occupational category and decade works similar to before. In other words, I am interested in the linear combination of the coefficients of the various interaction terms and the triple interaction. The calculation of the point estimates works similar to before, and for meaningful standard errors I now apply the formula for triple interactions by Aiken et al. (1991). Based on the model prediction in Proposition 1 is correct, then the u-shaped pattern of intergenerational elasticity across occupational backgrounds should either evolve or become stronger over time.

Figure 4 shows the same coefficient plots as the left panel in Figure 3 for the all decades since the 1980s. Panel (a) shows IGE by parental occupation during the 1980s, where a declining pattern across occupational group is recognizable, however there is no statistically significant difference between any pair of point estimates. During the 1990s, there is no sizeable nor statistical difference between all three groups in terms of the intergenerational mobility. In the 2000s, i.e. after the preceding boom in computer capital investment, the u-shaped pattern starts to emerge. Finally, in the 2010s when labor market polarization has progressed the most within the sample period, there is a clear u-shaped pattern across occupational groups. In the 2010s the point estimate of the IGE for children whose parents work in abstract occupations is even higher than for those children whose parents work in manual employment.

Overall, the strongest movements in intergenerational elasticity over time occur for children with parents working in either routine or abstract occupations. In particular, Figure 4 shows that the importance of parental income decreases the strongest over time for children with a routine background, whereas it increases significantly for children with an abstract background. The point estimate for the former falls from around .27 in the 1990s to below .15 in the 2010s, while the latter increases from .21 in the 1980s to close to .30 in the 2010s. On the other hand, the point estimate of intergenerational elasticity for children with parents working in manual employment stays largely constant with a value of .27. Only in the 2010s, there is a slight decrease to around .25, but the point estimate is not statistically significant from the previous decades.

Figure 4: IGE by Occupational Background and Decade



*Notes:* The four panels repeat the same analysis of intergenerational elasticity by parental occupation as in the left panel of Figure 3, but differentiates by decade. This allows for an analysis how IGE changed over time. It shows that the u-shaped pattern seen above occurs over time, in particular between the 1990s and 2000s, and is reinforced in the 2010s.

#### 4.2.3 Transition Matrices

Table 5 presents descriptive evidence for the transition between occupational groups and how it changes over time. Given that the price of capital fell over time in the last decades, which lead to an automation of routine occupations, the model predicts rising transitions from routine to both manual and abstract, and less transitions vice versa. Starting with Table 5a, which shows transitions between parental and child occupation between the three occupational groups of interest during the 1980s, it is evident that the majority of children is working in the same occupational group as their parents. This observation is not surprising given the evidence of strong intergenerational transmission of occupations between parents and children, shown, among others, by Hellerstein and Morrill (2011) and Ferrie (2005) for the United States. Over time, the “immobility” across occupational groups becomes stronger for both manual and abstract levels, whereas it falls for the case where parents work in routine occupations. Instead, the probabil-

Table 5: Transition Matrices between 1980 and 2017

(a) 1980s				(b) 1990s			
Parental Occupation	Child Occupation			Parental Occupation	Child Occupation		
	Manual	Routine	Abstract		Manual	Routine	Abstract
	%	%	%		%	%	%
Manual	41.5	29.5	29.0	Manual	43.5	26.7	29.8
Routine	28.8	36.7	34.4	Routine	35.4	34.0	30.6
Abstract	24.0	24.7	51.3	Abstract	21.5	22.9	55.6

(c) 2000s				(d) 2010s			
Parental Occupation	Child Occupation			Parental Occupation	Child Occupation		
	Manual	Routine	Abstract		Manual	Routine	Abstract
	%	%	%		%	%	%
Manual	51.3	21.1	27.6	Manual	48.1	22.1	29.9
Routine	33.9	25.3	40.7	Routine	39.2	22.8	38.0
Abstract	25.8	22.6	51.7	Abstract	27.7	19.5	52.8

*Notes:* The tables show the transition matrices between the three major occupational categories between parents and children. They show transition probabilities over time, that is by decade. On the left axis, the table lists the parental occupations, while it lists the same occupational groups for children at the horizontal axis.

ity to transition into either manual or abstract work is already higher than to stay in a routine occupation in the 2000s (see Table 5c).

### 4.3 Polarization of Educational Attainment

Educational choice is the mechanism how labor market polarization detrimentally affects upward mobility of children from low-income parents. The model predicts that falling prices for capital reduce the returns to secondary education, and therefore individuals are increasingly choose more extreme levels of education, i.e. primary or tertiary education. This section provides evidence that education has become more polarized in the United States over time, and that the increase in educational polarization is linked to local patterns in labor market polarization.

In order to quantify the impact of education, I estimate the previous model in equation (18 including covariates for both college and high-school education. Second, I estimate equation 20

excluding educational attainment in the vector of demographic characteristics. I compare the point estimates, where the differences between the respective coefficients allow me to quantify the impact of education on upward mobility. When including the principal components for both COL and K12 in the estimation of upward mobility, I obtain a point estimate of .38. Compared to the baseline estimate of .57, it falls by around 33%. The coefficients for intergenerational elasticity of income for all three categories are substantially higher. For parents working in manual, routine and abstract occupations, they are .387, .281 and .416, respectively. Compared to the point estimates including the individual level of education, the point estimates fall between 25 and 36 percent. Both these exercises provide evidence that the education accounts for at least 30 percent of intergenerational mobility.

I illustrate rising polarization in educational attainment in four stages. First, I calculate polarization indices for educational attainment of young labor market entrants for the United States as a whole. Second, I compute family premia based on Checchi et al. (2013) using the PSID. This allows me to compute how the probability of achieving a specific level of educational attainment depends on parental education for birth cohorts from 1955 until 1990. Third, I relate educational polarization with decadal changes in routine employment exploiting variation across time and commuting zones. Last, I focus on individual education decisions using data from the educational supplement of the Current Population Survey, and how they depend on both past regional polarization and family background.

#### *4.3.1 National Educational Polarization*

The model predicts an adjustment of labor supply with respect to education over generations as routine wages are declining. To see whether this holds true, I compute polarization indices for educational attainment using the Decennial Census for the years 1970 to 2000 and the American Community Service 2010 and 2018. I focus on respondents between the age of 20 and 29 as they have entered the labor market recently and respond the strongest to labor demand. Adão et al. (2020) provide evidence that changes in skill demand for cognitive skills have affected younger workers more than older workers. Further, I limit the sample to the labor force, i.e. I only exploit education decisions of employed or unemployed, and not individuals still in education or outside of the labor force for other reasons. Finally, I have excluded observations with zero or non-available schooling.<sup>15</sup>

Importantly, educational attainment is measured as a categorical variable in the Decennial Censuses and the ACS. Therefore, standard measures of spread or tailedness like standard deviation and kurtosis obsolete because they employ the mean as a centre in the space of the distribution. The same holds for the polarization index by Esteban and Ray (1994) as they consider the mean as the central point. Allison and Foster (2004) prove that the mean is not a suitable central point to measure polarization for categorical variables. Instead, the authors argue that median should

---

<sup>15</sup>In the Census files, the same value represents both possibilities. As it seems more likely to have no data on education than to observe individuals without any schooling, I decide to exclude these observations.

serve as the central point of the distribution. However, one issue with all measures which compute the polarization of distributions with ordinal data is that they require the same median in order to compare the polarization index.

Educational attainment is measured in eleven categories, and the values six and ten indicate high-school graduation and four years of college, respectively. All polarization indices are equal to zero if all individuals acquire the same level of education, i.e. when every labor market entrant has the same educational attainment equivalent to the median. The indices can rise for two reasons: first, the larger the share of young labor market entrants who do not acquire the same level as the median individual, and second, the further away the same share of individuals are from the median individual. In the case of education, the indices would reach their maximum if 50% of the individuals achieved the lowest level of educational attainment (grade 4) and the other 50% achieved the highest level of educational attainment (5+ years of college), with the median individual obtaining a high-school degree.

I compute various polarization indices for education. The first measure is based on Allison and Foster (2004), subsequently abbreviated “AF”. Its main drawback is its scale dependence. In other words, the maximum value of the AF index depends on the range of categories of the distribution. The second polarization index is based on Naga and Yalcin (2008), abbreviated by “ANY” in the subsequent tables. It has two advantages compared to the AF index. First, it is not scale dependent, instead it ranges between zero and one. Second, it allows me to put different weights on the mass below or above the median. This feature allows me to determine whether only one side of the distribution is affecting the overall index of educational polarization. Especially with respect to education and the upskilling over the last decades this feature provides useful insights. The index relies on two parameters, expressed as  $ANY(\alpha, \beta)$ . A higher value of  $\alpha$  indicates that more weight is put on the upper part of the distribution, i.e. above the median. Conversely, a higher value of  $\beta$  indicates that more weight is put on the lower part of the distribution, i.e. below the median. This means that I can determine with this index whether polarization is driven by movements above or below the median, or both. The baseline measure of  $ANY(1,1)$  is equal to AF divided by the number of categories.

Table 6 shows the development of educational polarization in the United States. Columns (1) and (2) indicate the census year used for the computation and which cohorts are included in the sample due to the restrictions explained above. Column (3) shows the median because all measures are median-preserving, i.e. the polarization over different distributions can only be compared if the median coincides between different distributions. The median labor market entrant from the cohorts 1941 to 1960 has a high-school degree. This changes for all subsequent cohorts, where the median individual acquires a level of education equivalent to one year of college. This allows me to compare young labor market entrants who were particularly affected by the rise of ICT capital.

Columns (4) to (8) present the different indices of the educational polarization for labor market entrants between 20 to 29. The table shows a strong increase in the AF indicated by



Table 6: Educational Polarization

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Year	Cohorts	Median	AF	ANY (2,1)	ANY (1,2)	ANY (4,1)	ANY (1,4)
1970	1941-50	6	3.15	.30	.37	.37	.47
1980	1951-60	6	3.11	.34	.38	.44	.47
1990	1961-70	7	2.72	.26	.31	.30	.39
2000	1971-80	7	2.99	.29	.35	.34	.43
2010	1981-90	7	2.96	.30	.35	.37	.43
2018	1989-98	7	3.01	.34	.37	.44	.45

*Notes:* This table shows various measures of educational polarization for young labor market entrants between the census years 1970 and 2018. It is important to keep in mind that all polarization indices are only comparable if the median in Column (3) is the same. Column (4) shows the Allison and Foster (2004) polarization measure (AF) and ANY( $\alpha$ ,  $\beta$ ) in Columns (5) to (8) measure polarization based on Naga and Yalcin (2008) with varying parameter values to put more weight on different parts of the distribution. Columns (5) and (7) put more weight on the upper part of the distribution, and columns (6) and (8) put more weight on the lower part of the distribution of educational attainment.

Column (4) between 1990 and 2000, i.e. during the time of the boom in computer capital. This view is confirmed by the ANY indices with varying parameters. Independent on whether I stress the upper part - reflected in Columns (5) and (7) by the higher value of  $\alpha$  - or the lower part - reflected in Columns (6) and (8) by the higher value of  $\beta$  - of the distribution, there is a substantial increase during the same period. Specifically, the rise in all ANY( $\alpha$ ,  $\beta$ ) indices, is about 10% over this period, which is equivalent to the overall change in AF in Column (4). After this strong rise, the polarization indices continue to increase slightly, but at a slower rate. The upper part of the education distribution constitutes an exception as indicated by Columns(5) and (7), it sees a strong rise between 2010 and 2018 as well.

Educational attainment of young labor force participants has become more polarized in the United States between 1990 and 2018. However, the polarization do not allow me to investigate whether individuals make extreme education choice influenced by labor market polarization nor family background.

#### 4.3.2 Family Premia

The main advantage of the Panel Study of Income Dynamics is that I can investigate educational choices by parental background. It is crucial to clearly identify the impact of parents on educational choices of young young labor market entrants. The standard approach typically estimates intergenerational transmission of education the following way:

$$c_i = \alpha + \rho f_i + \varepsilon_i,$$

where  $c_i$  and  $f_i$  are the number of years of education of child and father, respectively, from family  $i$ , each normalized by their respective standard deviation. The decomposition of  $\rho$  by Checchi et al. (2013) shows that the coefficient also reflects the general upskilling of the population during the second half of the last century. Based on the decomposition, the authors

propose multiple indicators of intergenerational persistence of education. I focus on the “family premium” for a given education level, which indicates whether a child experiences a benefit or a penalty of achieving the education level. Formally, it is defined as:

$$Pr(c = t|f = j) - Pr(c = t), \quad (21)$$

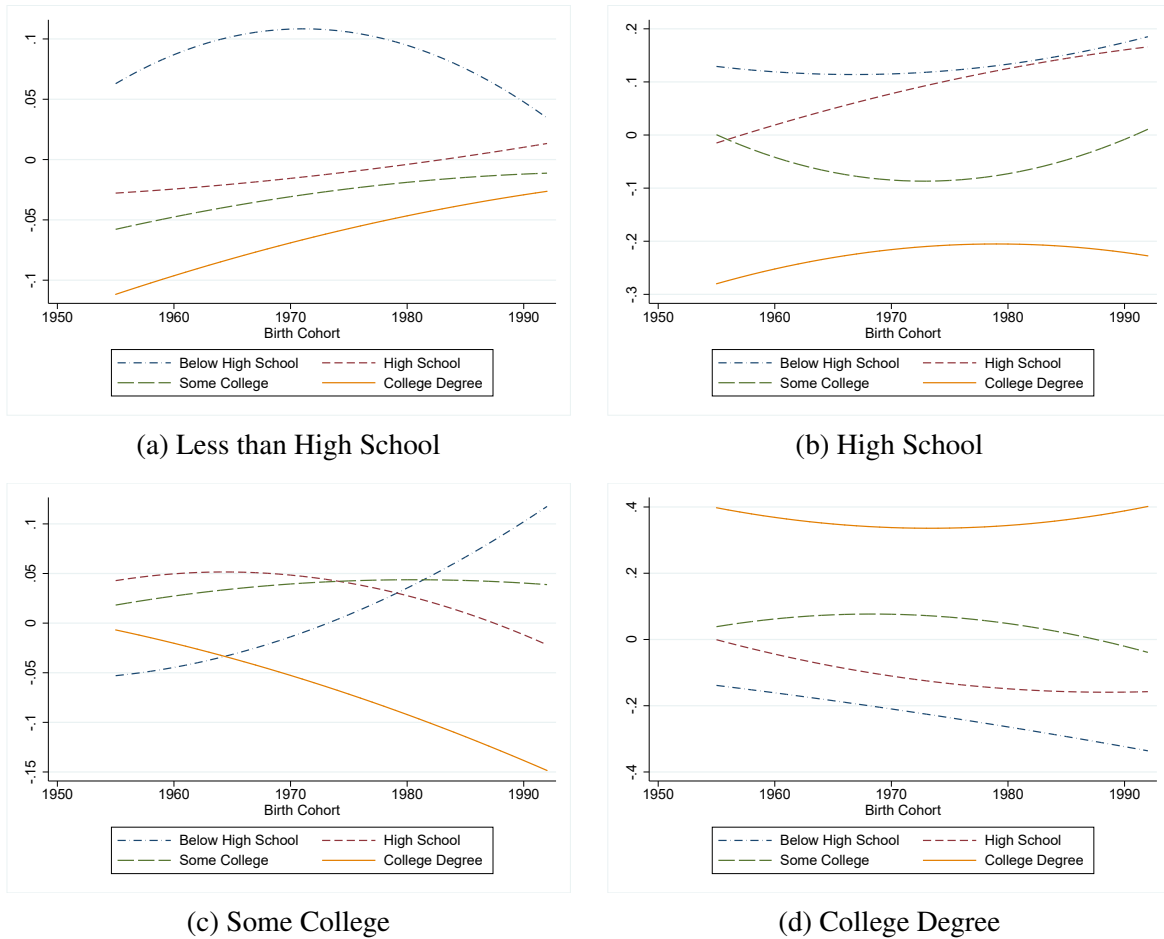
where  $t, j$  are specifying different education levels, which may coincide or not. I choose four levels, namely high school dropouts, high school graduates, some college and college degree.

Figure 5 show the results for the different family premia for a given education level  $t$  by background  $j$ . Panel (a) shows the family premia for dropping out of high school, for which all family premia are mainly rising, only for children whose parents also have less than high school education, the family premia premium is falling since the 1970 cohort. Panel (b) depicts the evolution of the family premia for high school graduation for all backgrounds. Interestingly, the family premia are rising for all backgrounds, albeit it rises the strongest for children whose parents are also high school graduates. The family premia for high school graduation for children whose parents have either some college or a college degree behave somewhat differently over time. The former first falls over time until the mid 1970s cohorts, and then rises again steeply, with a similar family premium for the 1955 and 1991 cohorts. The family premium where parents have a college degree rises until the 1980s cohort and then slightly declines.

Focusing on the two lower panels in Figure 5, the family premia for children whose parents are either high school or college graduates is falling over time. Treating “some college” as the category suitable for routine occupations, then the empirical evidence in panel (c) is largely in line with the model predictions: less children whose parents have “extreme” educational backgrounds tend to enter less the middle category of “some college”. On the other hand, the family premium is slightly rising, but is plateauing since the early 1970s cohort. The model would have predicted a decline for this development as well, but the family premium has to be considered in comparison to other backgrounds. One strong exception is the spectacular rise of the family premium for “some college” for children whose parents do not graduate high school. It is growing over time for all birth cohorts under consideration, and starts out with the lowest value, and in most of the 1980s it is already the highest.

Finally, panel (d) shows the family premia for a college degree depending on parental education. The first observation is that there is a strong level difference between family premia, in particular between those with parental college background and all other education levels. Further, over time this gap seems to widen. In other words, children in the 1980s experience a stronger family penalty when trying to obtain a college degree if their parents do not have a college degree themselves. Further, the family premium for children with parents holding a college degree is rising slightly over time after taking a dip until the 1970s birth cohorts. The latter is a clear prediction of the model as it indicates stronger “stickiness” for the upper tail of the education distribution. However, the model also predicts higher transition from secondary

Figure 5: Family Premia for Education Levels



*Notes:* The figure shows family premia dependent on educational background of the head for the PSID sample. The four panels present the family premia for a given level of educational attainment. Calculations of these premia are based on Checchi et al. (2013). Panel (a) depicts the family premia for less than high school dropouts, panel (b) for high school graduation, panel (c) for some college, and panel (d) for college graduation.

to tertiary education. This holds true until the 1970s cohorts, but the family premium has been declining since. This lack of increasing chances to obtain a college degree if the parents do not possess one by themselves can have severe consequences in terms of upward mobility as Altonji and Zhong (2020) illustrate how college degrees are crucial in terms of income.

Overall, the findings on educational polarization mirror the predictions from the model laid out in Section 2. It becomes increasingly difficult for children from parents without a college degree to obtain one themselves as shown by the family premia, based on Checchi et al. (2013), for this degree, while the chances have improved for children whose parents are college graduates. Investigating educational polarization with the family of indices by Naga and Yalcin (2008) also confirms most predictions from the model. First, polarization is falling for children whose parents are college graduates, indicating that children with this educational background tend to have less degrees outside of college degrees as well. While there is no indication of

rising educational polarization for children from less educated parental backgrounds with equal weights across the whole distribution, polarization is rising when either the lower or the upper tail are more weighted. Lastly, the latter is rising stronger, in line with empirical evidence, e.g. Autor and Dorn (2013), that wages for high-skilled workers rose stronger than those wages for low-skilled workers.

### 4.3.3 Regional Educational Polarization

This section presents empirical evidence that the local decrease in routine jobs leads raises the probability of young labor force participants to make more “extreme” educational choices. However, the advantage of adding geographic variation comes at the expense of not being able to retrieve information about the parental background. I compute the five indices of educational polarization from Table 6 for all commuting zones for the census years between 1970 and 2018. To determine the effect of labor market polarization, I estimate a similar model as for upward mobility in equation (18). The estimation equation takes on the following form:

$$EducPolarization_{jt} = \beta_0 + \beta_1 \Delta Routine_{jt-1} + X_{jt-1} \Theta + \gamma_t + \gamma_s + \varepsilon_{jt}, \quad (22)$$

where  $EducPolarization_{jt}$  denotes educational polarization indices of labor force participants between 20 and 29 in commuting-zone  $j$  in census year  $t$ . Educational polarization is explained by the change in local routine employment during the previous decade, i.e. before the individuals whose educational choice I measure actually enters the labor market. I include a vector of controls (also expressed as 10-year changes) based on the control variables by Autor and Dorn (2013).<sup>16</sup>

The estimation accounts for the median-preserving property of the polarization indices in two ways. First, I limit the sample to commuting-zones where median education is equal to the national mode of median educational across commuting-zones in a given year. I choose the national mode over the national median because it maximizes the number of commuting-zones in the sample. Second, I include time-fixed effects ( $\gamma_t$ ), meaning that I only exploit within-period variation when determining the impact of the decline in routine employment on educational polarization. Further, state-fixed effects ( $\gamma_s$ ) capture differences in the institutional framework, e.g. the age of compulsory schooling, ranging from 16 to 18. This can have implications on high-school dropout rates, which would imply stronger educational polarization.

The parameter of interest in this regression is  $\beta_1$ , which, according to the model, should be *negative* as a weaker decline in routine occupations implies less extreme education decisions. For comparability across indices and better interpretation of the coefficients, I normalize both the polarization indices and the change in routine employment.

---

<sup>16</sup>I exclude the ratio of college to non-college population as this educational polarization is a similar measure to this ratio, but focuses on the young labor force participants.

Table 7: Education Polarization and Changes in Routine Employment

	(1)	(2)	(3)	(4)	(5)
	AF	ANY(2,1)	ANY(1,2)	ANY(4,1)	ANY(1,4)
Panel A - OLS - Covariates as Differences					
$\Delta$ Routine	-0.28*** (-7.70)	-0.44*** (-10.41)	-0.40*** (-10.01)	-0.48*** (-10.16)	-0.42*** (-10.50)
R <sup>2</sup>	0.65	0.50	0.58	0.45	0.54
Panel B - OLS - Covariates as Lagged Levels					
$\Delta$ Routine	-0.08*** (-3.29)	-0.13*** (-4.95)	-0.12*** (-4.64)	-0.13*** (-4.85)	-0.12*** (-4.89)
Obs.	3543	3543	3543	3543	3543
F-Statistic					

*Notes:* This table estimates the impact of changes in routine employment per working-age population on various indices of educational polarization on the commuting-zone level for the workforce between 20 and 29. Commuting zones are chosen such that their median is equal to the national median in order to maximize the number of observations. Panel A shows the OLS results measuring the covariates as changes, and Panel B shows the OLS results measuring the covariates in (lagged) levels. Year- and state-fixed effects are not shown. Observations are weighted by population share in 1970. p-values are shown in brackets. \* denotes 10% significance, \*\* denotes 5% significance, \*\*\* denotes 1% significance.

Table 7 shows the results based on estimation equation (22) with the five different measures of educational polarization in columns (1) to (5). The panels differ in their estimation strategy, the upper one shows the OLS results, and the lower shows the results for the IV estimation strategy.

Both panels show that there is a negative impact of the 10-year change in routine employment on educational polarization of young labor force participants in the OLS estimation. In other words, the weaker labor market polarization in a commuting zone, the less polarized are education outcomes of young labor force entrants. In magnitudes, the point estimate in the first column in panel B a one-standard deviation decrease in routine employment raises the polarization index by 1.09 standard deviations. The coefficients are negative for all indices. Surprisingly, they are smaller for both ANY indices stressing the upper part of the distribution, i.e. ANY(2,1) and ANY(4,1) in columns (2) and (4), respectively. The difference in the point estimates indicates that young labor force participants are more likely to choose lower education levels in response to stronger labor market polarization. They also choose higher levels of educational attainment, albeit a lower rate.

Panel B shows the OLS estimates, where I use the covariates are expressed as lagged levels. Compared to Panel A, the point estimates are smaller in size, around the factor of 3.5. A one percentage point decrease in routine employment over 10 years increases educational polarization at least 10 percent of a standard deviation for all polarization indices.

#### 4.3.4 Individual Education Choices

In the final part of my analysis of educational decisions, I consider individual level data on education from the Current Population Survey (CPS) and merge it with state-specific declines in routine employment from the Decennial Censuses. This approach is equivalent to Ferriere et al. (2020), who investigate individual education decisions in the case of detrimental labor market shocks in the form of Chinese import competition. I estimate the linear probability model

$$e_{ist} = \sum_{q=1}^4 \beta_q \mathbb{1}_{\{Y_{irt} \in q\}} \Delta Routine_{st} + \delta_e \sum_q \bar{e}_{qst-1} + \gamma_s + \gamma_t + \epsilon_{ist}, \quad (23)$$

where  $e_{ist}$  denotes the individual education decision of individual  $i$  in state  $s$  at time  $t$ . I focus on four education decisions, namely high-school dropout, high-school graduation, some college and college. College combines both college enrolment and college degree.<sup>17</sup>  $\mathbb{1}_{\{Y_{irt} \in q\}}$  denotes an indicator function with value 1 whenever individual  $i$ 's household income is in quartile  $q$  of the overall income distribution. In order to account for peer effects,  $\bar{e}_{qst-1}$  measures the fraction of each income quartile with the same educational decision at  $t - 1$ .  $\gamma_s$  and  $\gamma_t$  indicate state- and time-fixed effects.

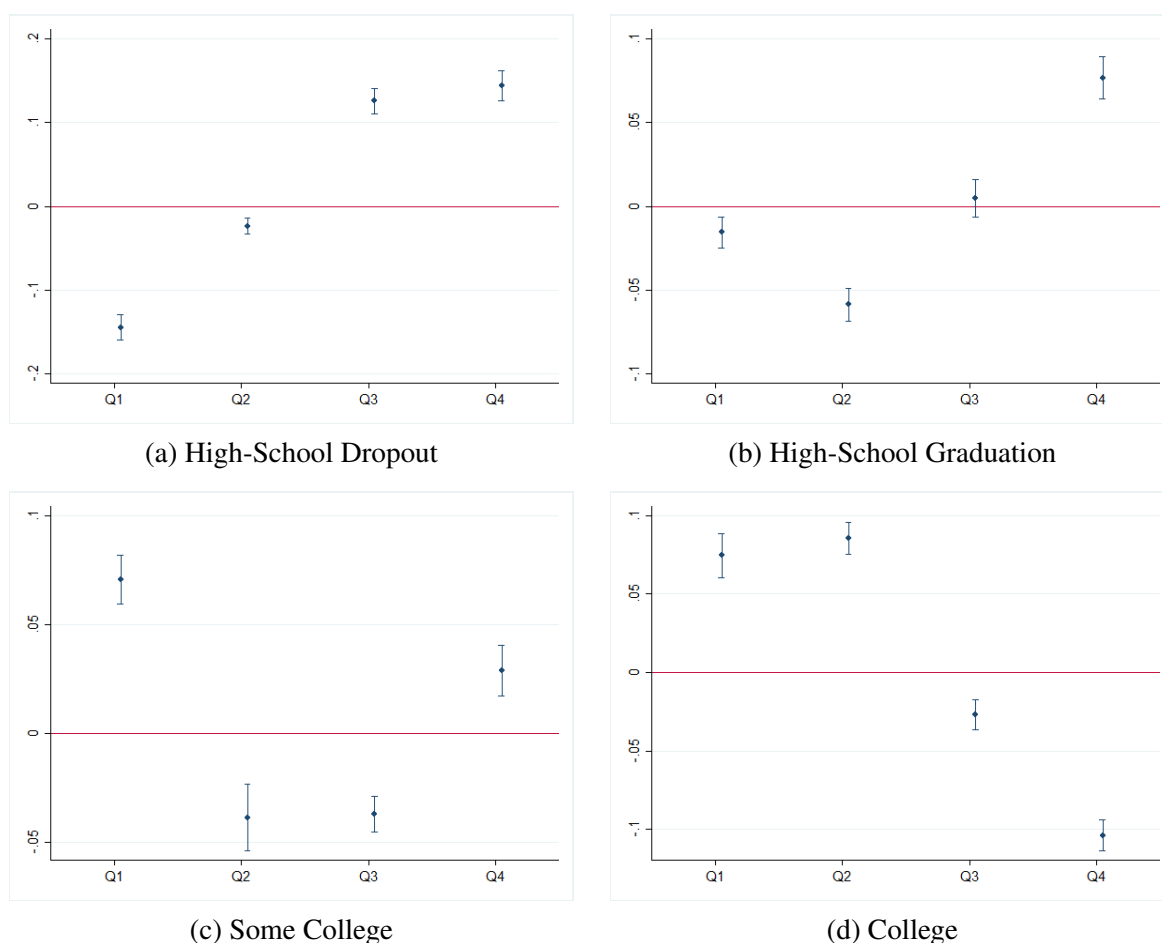
The parameters of interest are  $\beta_1$  to  $\beta_4$ . Each  $\beta_q$  provides information how the decline in routine employment in the previous decade affects the educational decision over all individuals conditional on the individual located in income quartile  $q$ . I expect the  $\beta_q$  point estimates to indicate that extreme educational choices are less probable if labor market polarization is weak. For example, I expect  $\beta_1$  to be negative if the dependent variable is high-school drop out. A negative sign would imply that children from families located in the first quartile of the income distribution are less likely to drop out of high-school if labor market polarization is weak. Analogously, I expect  $\beta_4$  to be negative if the college is the outcome variable.

Figure 6 shows the coefficient estimates of the linear probability model with four different educational choices. The coefficient plots allow for two comparisons. First, each graph compares how labor market polarization impacts the likelihood across children from all income quartiles ( $Q1$  to  $Q4$ ) to choose a given level educational attainment. The second comparison is to compare the point estimates for children from a given quartile and how labor market polarization affects all education choices.

The coefficient plots reveal that the labor market polarization has a similar impact on most education outcomes of the lower two quartiles. The upper two panels provide evidence that children who grow up in families below median income are less likely to choose either dropping out or finishing high-school if labor market polarization is weak. In other words, if the decline in routine-income jobs in the state where children is small before children enter the labor market, the children from parents with below-median income are less likely to choose lower levels of education, specifically to drop out or graduate from high-school. Instead, they are more likely to

<sup>17</sup>The individual graphs are qualitatively unchanged compared to the graph combining both education outcomes.

Figure 6: Individual Education Choices



The coefficient plots show the point estimates of the linear probability model in equation (23) for various education choices. In the top left panel the dependent variable is a binary indicator of high-school dropout, and in the top right panel the dependent variable is a binary variable of high-school graduation. The bottom panels focus on college education, with the left one on some college education, and in the right panel the education decision relates to college enrolment and degree.

choose higher education levels, albeit one important difference exists. Children from the lowest quartile are more likely to attain some college education, whereas children from the second quartile are more likely to be enrolled in college or have a college degree.

Labor market polarization also impacts the education choices of children from families above median income. Comparing all four panels, children from more affluent parents are less likely to choose the highest level of education if labor market polarization is weak, i.e. if the decline in routine employment is small. Surprisingly, they are more likely to drop out of high-school, and in the case of children from the highest income quartile, also to graduate from high-school or to obtain some college education if labor market polarization is weak. The strength of the decline in routine employment does not have a significant impact on children from the third quartile to obtain high-school education.

Overall, the results of the linear probability model indicate stronger “education stickiness” if

labor market polarization is high. With a strong decline in routine employment in the state where children live in before they enter the labor market, the probabilities to choose extreme educational attainment levels more similar to their parental background increases.

#### 4.4 Qualifying Predictions

This section confronts the remaining predictions with the data. First, the model also makes predictions how locations differ with respect to housing prices, demand for capital and thick tails. Specifically, the model predicts that high-productivity locations have higher populations and housing prices, demand more capital and exhibit thick tails. With the additional assumption of same area across locations, higher population also translates into higher density. In a second step I confirm that high-income parents provide more support to their children, both in absolute and relative terms. Therefore, the assumption of financial support proportional to parental income is conservative. Relaxing this assumption would reinforce the model predictions obtained in Section 2.

##### 4.4.1 ICT Investment, Housing Prices and Thick Tails

First, I investigate the spatial pattern of demand for computer capital with respect to log density. One of the key model predictions relates to productivity and density. The more productive location has a higher population size, and with the assumption of the same area, is also more densely population. Therefore, I use log density to represent productive regions.

The model predicts that locations with higher density demand more capital due to larger absolute number of workers in routine occupations. If the relationship between capital demand and historical density is positive, then historical density serves as a good predictor for the rise of ICT capital. Capital also performs routine tasks, and therefore replaces routine workers and makes investment in secondary education less profitable.

I measure ICT investment per worker, which is more forward-looking compared to actual computer adoption or IT budget per worker. Other studies, such as Beaudry et al. (2010) and Autor and Dorn (2013), use computer adoption, though there is likely a strong positive correlation between ICT investment and computer adoption. Eeckhout et al. (2019) use establishment-level data at the MSA (“Metropolitan Statistical Area”) in order to measure IT budget per worker. In order to measure ICT investment per worker for each commuting-zone, I apportion ICT investment to the region according to its share of national industry employment, i.e. shift-share measure:

$$ICT_{jt}^{PW} = \sum_k \frac{L_{kjt}}{L_{kt}} \frac{ICT_{kt}}{L_{jt}}. \quad (24)$$

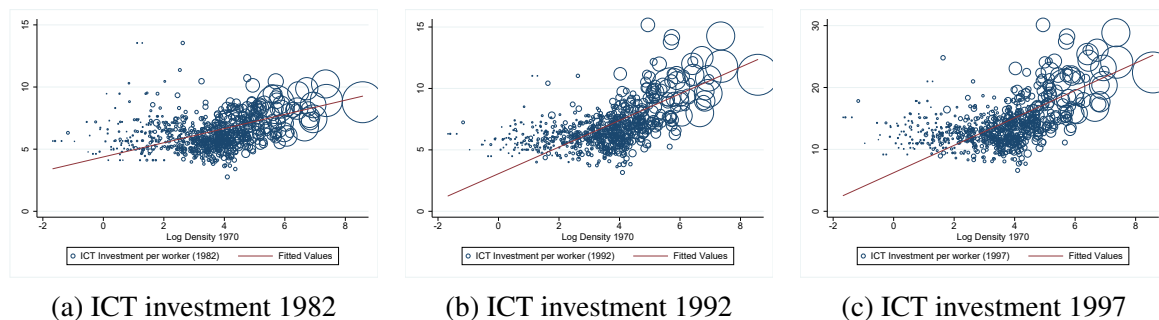
In this equation,  $L_{jt}$  measures total employment in location  $j$  in year  $t$ , and  $ICT_{kt}$  denotes total ICT investment in industry  $k$  in year  $t$ . The first fraction is a weighting factor, which exploits variation in the local employment structure. The variation arises mainly from the differences in



ICT-using industries across commuting zones.

Figure 7 measures plots the relationship between historical log density and ICT investment per worker for three different years, namely in 1982, 1992 and 1997. For comparability across time, I measure ICT investment per worker in 2010 USD. It is visible that ICT investment increases over time, in line with the IT boom in the 1990s. The unit of observation is the commuting zone in the United States. Observations are weighted by the share of population relative to the whole population in the United States in 1970.

Figure 7: Density and ICT Investment

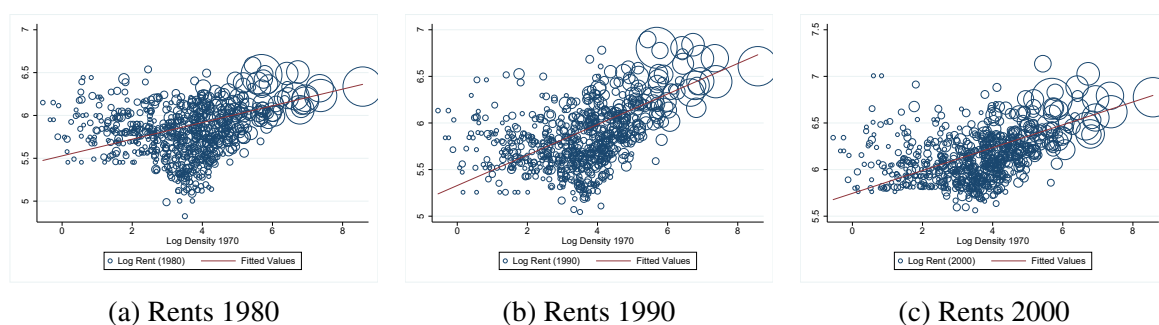


The scatter plots show the correlations between historical levels of density (expressed by its log) and ICT investment per worker in various years on the commuting-zone level. As expected from the model, the slope is positive. Observations are weighted by their 1980 population share.

All three scatter plots show a positive and economically meaningful relationship between ICT investment and historical density. Over time, i.e. across the three scatter plots, the relationship between historical density and ICT investment per worker increases. In the left panel, the slope is equal to .57 and statistically significant at all conventional levels. This means that an increase of density by 1% is associated with an increase of 5.70 USD (in 2010 value) of ICT investment per worker. The slope coefficient nearly doubles comparing ICT investment in 1982 and 1992, and more than doubles again between 1992 and 1997. In total, it nearly increases by 4 times in real terms between 1982 and 1997. This is reflective of the strong investment boom in the late 1990s prior to the Dotcom bubble burst in 2001. The graphs show that this investment boom largely concentrates in areas with high density and high human capital. This is true especially given the surprisingly large  $R^2$  of around .5 in five out of the six correlations, with the exception of ICT investment and density in 1982 with a value of .34. The strength of the relationship is based on regressions applying the same weights and are shown in table B1 in Appendix B.

Another model prediction relates to housing prices, specifically the model predicts higher housing prices in the more productive location. This is because individuals are geographically mobile without relocation costs and therefore utility is equal for all individuals with a given educational level. This assumption leads to real wage equalization across locations. As equations (5a) to (5c) indicate, wages depend on location-specific productivity levels. Therefore, housing prices need to be higher in the more productive, and therefore larger, location.

Figure 8: Density and Rents



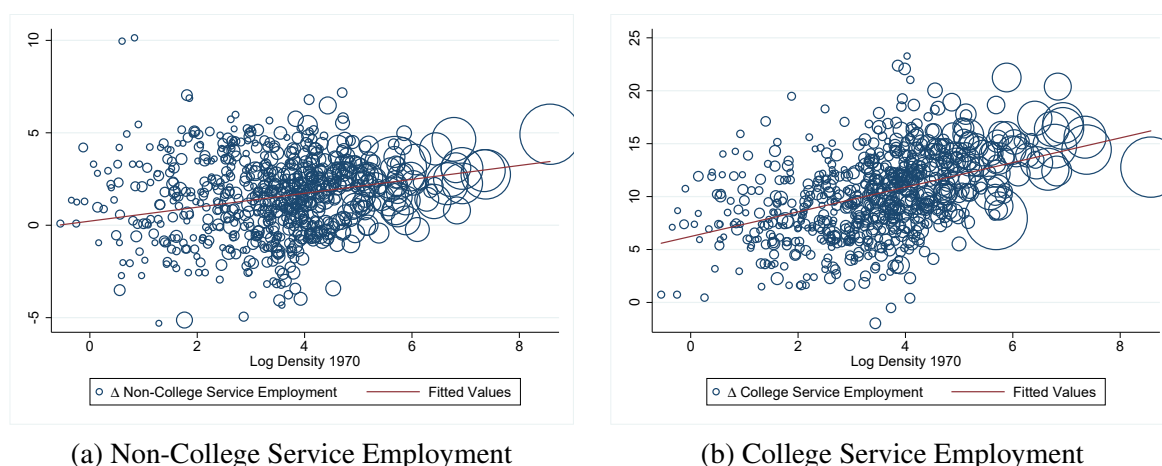
The scatter plots show the correlations between historical levels of density (expressed by its log) and rents in various years. As expected from the model, the slope is positive. Observations are weighted by their 1980 population share.

Figure 8 shows the relationship between average commuting-zone rental costs and historical (log) density for three different census years, where information on rents is available. In line with the model prediction, the slopes are positive in all three scatter plots. I show the relationship between rents and historical density because less low-income families own real estate, therefore rental costs are a better approximation for them. However, due to the strong relationship between local rental costs and local housing prices, the correlation between housing values and historical density is also positive. The scatter plots for the same years are shown in Figure B1 in Appendix B.

The last model prediction refers to thick tails. The rising capital stock reduces the returns to routine employment. As individuals decide on their educational attainment based on future wage ratios, they increasingly choose either primary or tertiary education. Subsequently, they work in either manual or abstract occupations, while the number of workers of routine tasks is shrinking. As shown above, high-productivity locations demand more capital, and therefore the share of manual and abstract workers should increase stronger in high-density commuting zones. To test this prediction, I relate the changes in college and non-college service employment to historical density.

Figure 9 shows that the tails of the occupational distribution are increasing stronger in high-density commuting zones. The left panel relates the change in non-college service (manual) employment between 1970 and 2010 with log density in 1970. In line with the model prediction, the slope is positive. The right panel plots the change in college service (abstract) employment against historical density. The slope is also positive, and stronger than for the change in non-college service employment. The model predicts equal changes across locations, but this is the result of the simplifying assumption of symmetry between manual and abstract employment.

Figure 9: Density and Thick Tails



The scatter plots show the correlations between historical levels of density (expressed by its log) and thick tails on the commuting-zone level. The left panel shows the change in non-college service employment between 1970 and 2010 on the vertical axis, while the right panel depicts change in college service employment between 1970 and 2010 on the vertical axis. As expected from the model, the slopes are positive. Observations are weighted by their 1980 population share.

#### 4.4.2 Intergenerational Transfers

I conjecture that parents bequeath the same proportion of their income to their children independent of their occupation, and hence income. Previous work on bequest motives across the income distribution, e.g. Menchik and David (1983) have established that richer parents tend to bequest more to their children. In the model, this assumption leads to more equal chances of children from low- and middle-income parents. In other words, it increases the chances of children whose parents work in manual or routine occupations to attain tertiary education and to experience upward mobility with respect to education, occupation and income.

In order to test the assumption, I make use of two supplemental studies from the PSID, i.e. the “Transfers Module” in 1988 and the “Rosters and Transfer Module” in 2013. They contain information on money given to children (in USD) and on how many hours parents help their children. I can link both supplemental studies to other information in the PSID and determine exact intergenerational transfers and how they differ by parental occupational group. Importantly, I investigate both absolute and proportional pecuniary support from parents to their children. I compute the proportional share of money support to their children relative to the definition of permanent income, i.e. I use parental income when children were between 13 and 19.

Tables 8 and 9 depict parental income, money support from parents in dollars and in percent, and time help. The latter also includes information on the share of parents provide pecuniary support to their children in school. Consistent with the ranking of wages in the model, average parental income is highest for abstract occupations, and followed by routine and manual occupations. Similarly, children get more money in absolute terms in this ranking in 1988, whereas there is a small difference in 2013 where children from parents working in manual occupations

receive more absolute money on average. Money support relative to permanent income during teenagehood does not indicate that the assumption is strongly violated, i.e. they are mainly in line with the previous literature: richer parents tend to spend more on their children.<sup>18</sup> The percentage in columns (3) is not equal to the proportion of money support in column (2) to parental income in column (1), but instead this reflects the average proportional change of the full sample available in each supplemental study.

Table 8: Parental Transfers to Children (1988)

Occupational Group	(1) Parental Income	(2) Money from Parents (USD)	(3) Money from Parents (%)	(4) Time Help from Parents
Manual	51458.7	249.26	.48	141.08
Routine	57737.93	529.89	.88	104.69
Abstract	80211.02	613.73	.81	72.37

*Notes:* This table shows parental income and transfers of time and money to their children. It shows that income is rising with occupational group, and so are monetary transfers by parents, both in absolute and percent. Help in terms of hours declines with occupational groups. The data is taken from the in 1988 by the PSID.

Columns (4) and (5) in Tables 8 and 9, respectively, show the average time support in hours for the full calendar year by parental occupation. Interestingly, this relationship across groups has completely flipped over time. In 1988, the relationship of money support (in USD) and time was inverse, that is the less money parents gave to their children for support, the more hours they helped their children. By 2013, this has changed, i.e. parents providing the most support in terms of money also provide the most support in terms of hours.

Table 9: Parental Transfers to Children (2013)

Occupational Group	(1) Parental Income	(2) Money from Parents (USD)	(3) Money from Parents (%)	(4) Support School (%)	(5) Time Help from Parents
Manual	57762.11	498.14	.74	15.54	77.12
Routine	63677.86	493.03	.82	19.74	80.69
Abstract	91012.91	1049.54	1.08	34.87	85.81

*Notes:* This table shows parental income and transfers of time and money to their children. It shows that income is rising with occupational group, and so are monetary transfers by parents, both in absolute and percent. Help in terms of hours declines with occupational groups. The data is taken from the in 2013 by the PSID.

Finally, Table 9 shows the share of parents supporting their children in school. The question in the supplemental PSID study asks *whether* parents support their child in school. A clear ranking in line with the wage ranking in the model and in the table is evident, namely a much

<sup>18</sup>There is a small discrepancy from these findings in relative support for children from routine parents relative to abstract parents in 1988, but the difference is not large.

larger share of parents with abstract occupations support their children in school compared to parents in manual and routine occupations. This is probably driven by stronger educational attainment for children from high-income parents and stronger financial support, and therefore already indicative of a sticky upper tail. The difference between support for educational attainment between routine and manual parents is not as large, but still sizeable.

#### 4.5 Robustness Analysis

Dauth (2014) argues that there is no good measure of labor market polarization. In the baseline estimations, I measure labor market polarization with changes in routine employment because it is the most remarkable consequence of routine-biased technological change. In line with Autor and Dorn (2013), I proxy labor market polarization using the rise in non-college service employment.

Table 10: Effect of labor market polarization on upward mobility

<b>I. 2<sup>nd</sup> Stage</b>	(1)	(2)	(3)
	OLS	OLS	IV
$\Delta$ Non-College Svc. Emp.	-0.92*** (-5.05)	-0.26*** (-3.02)	-2.17*** (-4.01)
State FE	No	Yes	Yes
Controls	No	Yes	Yes
Obs.	693	693	693
R <sup>2</sup>			
F-Statistic			20.73
<b>II. 1<sup>st</sup> Stage</b>			
Log Density (1970)			0.24*** (4.07)
Routine Emp. (1970)			0.04*** (3.38)
Partial R <sup>2</sup>			

*Notes:* The dependent variable of the second stage of the IV approach is absolute upward mobility regressed on labor market polarization measured as the change in non-college service employment between 1990 and 2010. The lower panel shows the first-stage results. Control variables and state fixed effects are not shown. Observations are weighted by population share in 1980. t-statistics are shown in brackets. \* denotes 10% significance, \*\* denotes 5% significance, \*\*\* denotes 1% significance.

Table 10 shows the results for equations (18) and (19) using an alternative proxy for labor market polarization. As this proxy is rising with labor market polarization (opposite to changes in routine employment), I expect the opposite sign of the estimated coefficient  $\beta_1$  compared to the benchmark results, i.e. the sign of the coefficient should be negative if labor market polarization has an adverse impact on intergenerational mobility. Comparing the OLS results in Table 10 with those in Table 3, the magnitudes are very similar, albeit the point estimates with changes in non-college service employment are slightly larger. Comparing columns (1) and (2)

of Tables 3 and 10, the magnitude of the point estimates in the latter are 19% and 13% higher than in the former.

The crucial difference between Tables 3 and 10 is visible in the instrumental variable estimation in column (3) of each table. The point estimate of the IV identification strategy is increasing (in magnitude) by much more in the robustness analysis compared to the benchmark estimation. While the OLS results are biased towards zero, the increase in the point estimates from column (2) to column (3) is unusual. The first-stage results do not indicate any weakness in terms the F-statistics or either the signs or statistical significance. The F-statistic is well above 10, and the point estimates of the first-stage in the lower panel indicate that the instruments, in line with expectations, raise the share of non-college service employment.

In order to confirm the results obtained above, I use two further measurements of intergenerational mobility taken from Chetty et al. (2014). The first is relative upward mobility, which measures the expected rank of children from the richest versus the poorest families. This means that it behaves opposite to absolute upward mobility, i.e. a higher value implies a lower degree of intergenerational mobility. Consequently, the expected sign for the estimated coefficient of  $\beta_1$  in equation (18) turns negative, i.e. the less strong labor market polarization, the smaller the distance between the expected ranks from families at the top and bottom of the income distribution. The second measure is the probability of an intergenerational transition between the lowest income quintile and the highest income quintile. This measure is, just like the benchmark results, from the perspective of children from low-income families, hence the parameter  $\beta_1$  should be positive (again).

Because the first-stage results the same as in Tables 3 and 4, Table 11 only presents the results of the OLS regression and for the second stage of the IV estimation strategy. For most specifications, the point estimate is - as expected - positive, indicating that in commuting zones with stronger labor market polarization kids from top-income parents experience higher expected ranks relative to children from low-income parents. As argued above for absolute mobility, endogeneity is also a concern in these estimation, leading to a downward bias for relative mobility (due to its inverse behavior) and an upward bias for the quintile transition probability. Using the point estimate in column (2) of panel A, a one percentage point increase in polarization (measured by a decline in routine employment) leads to an improvement of .53 ranks of children of the richest parents compared to children from the poorest. The point estimates in panel A indicate that children from the highest-ranked families profit more from labor market polarization than children from the lowest-ranked families. Precisely, the difference in percentiles of children from the lowest-ranked families and from the highest-ranked is equal to 4.09 percentiles comparing commuting zones without labor market polarization to the median-hit commuting zone. One reason why this is smaller than for the baseline results for children at the 25<sup>th</sup> percentile is that wages also changed non-monotonically. Autor and Dorn (2013) actually show that wage losses were the highest for parents at the 25<sup>th</sup> skill percentile, whereas workers below this skill percentile sometimes even experienced wage increases.

Table 11: Robustness analysis: Effect of labor market polarization on upward mobility

I. 2 <sup>nd</sup> Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	IV	MIG	LAB	COL	TAX	FAM	SOC	K12	INC	SEG	BLA
$\Delta$ Routine Emp.	-0.15*** (-2.59)	-0.53*** (-3.30)	-0.42*** (-2.65)	-0.50*** (-3.05)	-0.48*** (-2.85)	-0.57*** (-3.68)	-0.07 (-0.51)	-0.61*** (-3.71)	-0.44** (-2.09)	-0.61*** (-2.73)	0.10 (0.55)	-0.24 (-1.61)
Obs.	693	693	691	693	563	693	693	666	663	693	693	693
F-Statistic		122.32	113.67	125.91	110.47	114.48	104.76	100.44	78.24	61.21	46.59	105.68
Panel A - Relative Mobility												
$\Delta$ Routine Emp.	0.16*** (4.52)	0.44*** (4.44)	0.43*** (4.18)	0.44*** (4.37)	0.40*** (3.99)	0.18** (2.16)	0.48*** (4.48)	0.43*** (4.09)	0.37*** (2.60)	0.63*** (4.70)	-0.04 (-0.33)	0.36*** (3.78)
Obs.	693	693	691	693	563	693	693	666	663	693	693	693
F-Statistic		122.32	113.67	125.91	110.47	104.76	114.48	100.44	78.24	61.21	46.59	105.68
Panel B - Transition Probability (P1, K5)												

Notes: This table replicates the regression results from Tables 3 and 4, but exploits alternative measures of intergenerational mobility. Panel A uses relative mobility, while Panel B uses the probability of a child to be in the top quintile (K5), while parents were located in the lowest quintile (P1). Both measures are taken from Chetty et al. (2014). As before, MIG stands for migration, LAB for labor markets, COL for college, TAX for the tax system, FAM for family, SOC for social capital K12 for high-school, INC for properties of the income distribution, SEG for segregation, and BLA for fraction of black population. Control variables and state fixed effects are not shown. Observations are weighted by population share in 1980. t-statistics are shown in brackets. \* denotes 10% significance, \*\* denotes 5% significance, \*\*\* denotes 1% significance.

Panel B in Table 11 use the probability for a child reaching the highest quintile in the income distribution conditional on his parents being located in the lowest quintile show that labor market polarization also has a detrimental impact. Excluding the specification with the first principal component of segregation, the sign is negative in all specifications and statistically significant. Considering the point estimate of column (2), it shows that a one percentage point increase in routine employment reduces the transition probability by 0.44 percentage points. 3.42% more of the children from families in the lowest quintile of the income distribution in commuting zones without labor market polarization would reach the highest income quintile compared to children from the median-hit commuting zone. This difference is nearly equal to the difference between the 25<sup>th</sup> and the 75<sup>th</sup> of this mobility measure (3.85%).

## 5 Conclusion

Does labor market polarization limit the equality of opportunity for children from low-income parents? To understand the relationship between both phenomena, I set up a simple theoretical framework drawing on existing models on educational choice and labor market polarization. The model features substitutability between routine employment and capital and exogenous location-specific productivity differences. The education decision of the young generation depend on future wage ratios and parental bequests. Due to the substitutability between routine labor and capital, a rising capital stock reduces the returns to routine employment. Location-specific productivity differences impact local demand for capital, hence the decline in routine employment is stronger in high-productivity regions. The model also predicts that more productive regions are larger and more dense.

Subsequently, I take the model predictions to the data. The key model prediction relates to upward mobility of children from low-income parents. The decline of both routine wages and jobs reduces the chances of children to climb the economic ladder. I estimate the impact of labor market polarization on absolute upward mobility from Chetty et al. (2014) exploiting commuting-zone variation in the United States. To tackle various endogeneity issues I estimate the causal impact with an instrumental variable estimation strategy. Based on the model and the previous literature, I use (log) density and the historical share of routine employment before the arrival of the first commercially successful personal computers as an instrumental variable. The results provide evidence that labor market polarization substantially reduces the expected rank for children whose parents are located at lower parts of the income distribution.

Using intergenerational elasticity as another measure of intergenerational mobility, I show that the model prediction over time also holds in the data. With less intergenerational transitions out of manual and abstract occupations, the model predicts a rising IGE for these two groups. On the other hand, for children whose parents work in routine-intensive occupations, parental income the model predicts a lower importance of parental income as they transition into both manual and abstract occupations. I exploit information on children and their parents from the



PSID and test these model predictions. This pattern holds overall in the data, and the pattern evolves over time as labor market polarization progresses. However, the main changes in intergenerational elasticity take place for children from both routine and abstract occupations, whereas IGE stays at a similar level for children with parents in manual occupations.

The empirical evidence supports that the importance of the education channel via which labor market polarization reduces intergenerational mobility. First, I provide evidence that education of young labor market entrants has become more polarized over time in the United States. Second, I compute family premia for different levels of educational attainment, which suggest rising stickiness with respect to education. Third, I show that a decline in routine jobs raises local educational polarization for panel of commuting zones. The results suggest that labor market polarization drives extreme educational choices at both tails of the distribution. In the fourth and final step I estimate how labor market polarization affects individual educational choice depending on parental income. The findings support the notion of educational stickiness if labor market polarization is strong.

One concern with this paper is to ask whether the results also hold in countries where tuition fees are low or equal to zero. While the model assumes pecuniary bequests, Mayer et al. (2019) underline that other factors such as parental support and encouragement also matter for the decision on educational attainment. If non-pecuniary elements also matter for educational attainment, then the results of this study are likely to hold in other countries as well. For example, Landersø and Heckman (2017) argue that the mobility pattern between Denmark and the United States is “remarkably similar”, whereas Andrade and Thomsen (2018), using the same data, conclude that mobility is substantially higher in Denmark. The data on regional upward mobility for Italy by Güell et al. (2018) combined with data on regional labor market polarization can provide insights into whether the results change substantially as tuition fees in Italy are much lower than in the United States.

Overall, these findings are important for a variety of reasons. First, the findings in this paper contend with the notion that neighborhood is the main explanatory factor of intergenerational mobility as argued in Chetty and Hendren (2018a) and Chetty and Hendren (2018b). Instead, it highlights the role of labor markets as suggested by Rothstein (2019). Secondly, the paper highlights the role of structural transformations, subsequent wage declines for certain occupations, and how the next generation with parents who work in these vulnerable occupations is affected. Therefore, in order to allow children from all background to achieve their potential, it is crucial to identify vulnerable occupations during structural transformations, e.g. for the future the rise of robots as documented by Graetz and Michaels (2018) and Acemoglu and Restrepo (2020). Third, the findings indicate that segregation by race and/or income can also play a role, in line with findings by Fogli and Guerrieri (2019) and Chetty et al. (2020).

## References

- Abraham, A. (2008). Earnings inequality and skill-biased technological change with endogenous choice of education. *Journal of the European Economic Association*, 6(2-3):695–704.
- Abramitzky, R., Boustan, L. P., Jácome, E., and Pérez, S. (2019). Intergenerational mobility of immigrants in the us over two centuries. Technical report, National Bureau of Economic Research.
- Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: evidence from us local labor markets. *Journal of Political Economy*, (forthcoming).
- Adão, R., Beraja, M., and Pandalai-Nayar, N. (2020). Technological transitions with skill heterogeneity across generations. Technical report, National Bureau of Economic Research.
- Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., and Hémous, D. (2018). Innovation and top income inequality. *The Review of Economic Studies*, 86(1):1–45.
- Aiken, L. S., West, S. G., and Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.
- Allison, R. A. and Foster, J. E. (2004). Measuring health inequality using qualitative data. *Journal of health economics*, 23(3):505–524.
- Altonji, J. G. and Zhong, L. (2020). The labor market returns to advanced degrees. *Journal of Labor Economics* (forthcoming).
- Andrade, S. B. and Thomsen, J.-P. (2018). Intergenerational educational mobility in denmark and the united states. *Sociological Science*, 5:93–113.
- Andrews, R., Casey, M., Hardy, B. L., and Logan, T. D. (2017). Location matters: Historical racial segregation and intergenerational mobility. *Economics Letters*, 158:67–72.
- Autor, D. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–97.
- 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. (2019). Work of the past, work of the future. In *AEA Papers and Proceedings*, volume 109, pages 1–32.
- Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333.
- Beaudry, P., Doms, M., and Lewis, E. (2010). Should the personal computer be considered a technological revolution? evidence from us metropolitan areas. *Journal of Political Economy*, 118(5):988–1036.
- Black, S. and Devereux, P. J. (2011). Handbook of labor economics. ed. O. Ashenfelter and D. Card Vol. IVb, Chapter Recent Developments in Intergenerational Mobility, pages 1487–1541.

- Blanden, J., Goodman, A., Gregg, P., and Machin, S. (2004). Changes in intergenerational mobility in Britain. *Generational income mobility in North America and Europe*, pages 122–46.
- Boustan, L. P. (2013). Racial residential segregation in American cities. Technical report, National Bureau of Economic Research.
- Brambor, T., Clark, W. R., and Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political analysis*, 14(1):63–82.
- Castro, R. and Coen-Pirani, D. (2016). Explaining the evolution of educational attainment in the United States. *American Economic Journal: Macroeconomics*, 8(3):77–112.
- Checchi, D., Fiorio, C. V., and Leonardi, M. (2013). Intergenerational persistence of educational attainment in Italy. *Economics Letters*, 118(1):229–232.
- Chetty, R. and Hendren, N. (2018a). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.
- Chetty, R. and Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.
- Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2020). Race and economic opportunity in the United States: An intergenerational perspective. *The Quarterly Journal of Economics*, 135(2):711–783.
- 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.
- Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, 27(3):79–102.
- Dabla-Norris, M. E., Pizzinelli, C., and Rappaport, J. (2019). *Job Polarization and the Declining Fortunes of the Young: Evidence from the United Kingdom*. International Monetary Fund.
- Dauth, W. (2014). Job polarization on local labor markets. Technical report, IAB-Discussion Paper.
- Davis, D. R., Mengus, E., and Michalski, T. (2020). Labor market polarization and the great divergence: Theory and evidence. mimeo.
- Eeckhout, J., Hedtrich, C., Pinheiro, R., et al. (2019). Automation, spatial sorting, and job polarization. In *IZA 2019 Meeting Papers*. Society for Economic Dynamics.
- Eeckhout, J., Pinheiro, R., and Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, 122(3):554–620.
- Esteban, J.-M. and Ray, D. (1994). On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, pages 819–851.
- Fender, J. (2005). Self employment, education and credit constraints: A model of interdependent credit rationing decisions. *Journal of Macroeconomics*, 27(1):31–51.

- Ferrie, J. P. (2005). History lessons: The end of american exceptionalism? mobility in the united states since 1850. *Journal of Economic Perspectives*, 19(3):199–215.
- Ferriere, A., Navarro, G., Reyes-Heroles, R., et al. (2020). Escaping the losses from trade: The impact of heterogeneity on skill acquisition.
- Fogli, A. and Guerrieri, V. (2019). The end of the american dream? inequality and segregation in US cities. Technical report, National Bureau of Economic Research.
- Gallagher, R., Kaestner, R., and Persky, J. (2019). The geography of family differences and intergenerational mobility. *Journal of Economic Geography*, 19(3):589–618.
- Galor, O. and Tsiddon, D. (1997). Technological progress, mobility, and economic growth. *American Economic Review*, 87(3):363–382.
- Graetz, G. and Michaels, G. (2018). Robots at work. *Review of Economics and Statistics*, 100(5):753–768.
- Güell, M., Pellizzari, M., Pica, G., and Rodríguez Mora, J. V. (2018). Correlating social mobility and economic outcomes. *The Economic Journal*, 128(612):F353–F403.
- Harrison, B. and Bluestone, B. (1990). Wage polarisation in the us and the 'flexibility' debate. *Cambridge journal of Economics*, 14(3):351–373.
- Hellerstein, J. K. and Morrill, M. S. (2011). Dads and daughters the changing impact of fathers on women's occupational choices. *Journal of Human Resources*, 46(2):333–372.
- Hershbein, B. and Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, 108(7):1737–72.
- Hillman, N. W. (2016). Geography of college opportunity: The case of education deserts. *American Educational Research Journal*, 53(4):987–1021.
- Huggett, M., Ventura, G., and Yaron, A. (2011). Sources of lifetime inequality. *American Economic Review*, 101(7):2923–54.
- Jaimovich, N. and Siu, H. E. (2012). The trend is the cycle: Job polarization and jobless recoveries. Technical report, National Bureau of Economic Research.
- Kambourov, G. and Manovskii, I. (2009). Occupational mobility and wage inequality. *The Review of Economic Studies*, 76(2):731–759.
- Keane, M. P. and Wolpin, K. I. (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review*, 42(4):1051–1103.
- Kourtellos, A., Marr, C., and Tan, C. M. (2020). Local intergenerational mobility. *European Economic Review*, page 103460.
- Krueger, A. (2012). The rise and consequences of inequality. *Speech at the Center for American Progress, Washington, DC, January 12*.
- Landersø, R. and Heckman, J. J. (2017). The scandinavian fantasy: The sources of intergenerational mobility in denmark and the us. *The Scandinavian journal of economics*, 119(1):178–230.

- Lee, C.-I. and Solon, G. (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics*, 91(4):766–772.
- Major, L. E., Machin, S., et al. (2019). Social mobility. Technical report, Centre for Economic Performance, LSE.
- Manning, A. (2004). We can work it out: the impact of technological change on the demand for low-skill workers. *Scottish Journal of Political Economy*, 51(5):581–608.
- Maoz, Y. D. and Moav, O. (1999). Intergenerational mobility and the process of development. *The Economic Journal*, 109(458):677–697.
- Mayer, S. E., Kalil, A., Oreopoulos, P., and Gallegos, S. (2019). Using behavioral insights to increase parental engagement the parents and children together intervention. *Journal of Human Resources*, 54(4):900–925.
- Mazumder, B. (2018). Intergenerational mobility in the united states: What we have learned from the psid. *The Annals of the American Academy of Political and Social Science*, 680(1):213–234.
- Menchik, P. L. and David, M. (1983). Income distribution, lifetime savings, and bequests. *The American Economic Review*, 73(4):672–690.
- Naga, R. H. A. and Yalcin, T. (2008). Inequality measurement for ordered response health data. *Journal of Health Economics*, 27(6):1614–1625.
- Osterman, P. and Shulman, B. (2011). *Good Jobs America*. Russell Sage Foundation.
- Owen, A. L. and Weil, D. N. (1998). Intergenerational earnings mobility, inequality and growth. *Journal of Monetary Economics*, 41(1):71–104.
- Palomino, J. C., Marrero, G. A., and Rodríguez, J. G. (2018). One size doesn't fit all: A quantile analysis of intergenerational income mobility in the us (1980–2010). *The Journal of Economic Inequality*, 16(3):347–367.
- Papageorge, N. W. and Thom, K. (2020). Genes, education, and labor market outcomes: evidence from the health and retirement study. *Journal of the European Economic Association*, 18(3):1351–1399.
- Quillian, L. (2014). Does segregation create winners and losers? residential segregation and inequality in educational attainment. *Social Problems*, 61(3):402–426.
- Rothstein, J. (2019). Inequality of educational opportunity? Schools as mediators of the intergenerational transmission of income. *Journal of Labor Economics*, 37(S1):S85–S123.
- Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M. (2018). Ipums usa: Version 8.0 [dataset]. *Minneapolis, MN: IPUMS*.
- Sahota, G. S. (1978). Theories of personal income distribution: a survey.
- Sharkey, P. and Torrats-Espinosa, G. (2017). The effect of violent crime on economic mobility. *Journal of Urban Economics*, 102:22–33.

- Solon, G. (1992). Intergenerational income mobility in the united states. *The American Economic Review*, pages 393–408.
- Solon, G. (2002). Cross-country differences in intergenerational earnings mobility. *Journal of Economic Perspectives*, 16(3):59–66.
- Tan, H. (2019). A different land of opportunity: The geography of intergenerational mobility in the early 20th century us. Technical report.
- Tobin, J. (1982). *Asset accumulation and economic activity: Reflections on contemporary macroeconomic theory*. University of Chicago Press.
- vom Lehn, C. (2019). Labor market polarization, the decline of routine work, and technological change: A quantitative analysis. *Journal of Monetary Economics*.

## Appendix A

### A.1 Polarization over time

As this appendix concentrates on how rising capital stock in the economy affects educational choices over time, I drop the subscript  $j$  from the notation here. All derivations hold for all locations, independent of their difference in exogenous productivity  $A_j$ .

The exogenous driving force in the model is the technology parameter of capital  $\zeta_t$ . Given the assumed production cost of this intermediate good, it is equal to price of capital ( $p_{kt}$ ). The capital stock rises with falling technology parameter  $\zeta_t$ . Importantly, if the technology parameter is sufficiently large, there won't be any capital produced, and the economy produces the final good with three factors of production, i.e. manual, routine and abstract labor.

A rising capital stock in the economy detrimentally impact routine wages, but affect neither manual or abstract wages:

$$\frac{\partial w_{mjt}}{\partial K} = 0 \quad (25a)$$

$$\frac{\partial w_{rjt}}{\partial K} < 0 \quad (25b)$$

$$\frac{\partial w_{ajt}}{\partial K} = 0 \quad (25c)$$

Individuals take into account wage ratios when making their choice on education as seen in equations (11) to (16). The three wage ratios of importance react distinctly on a rise of capital (induced by falling price of capital):

$$\frac{\partial(w_{mjt}/w_{ajt})}{\partial K} = 0 \quad (26a)$$

$$\frac{\partial(w_{mjt}/w_{rjt})}{\partial K} > 0 \quad (26b)$$

$$\frac{\partial(w_{rjt}/w_{ajt})}{\partial K} < 0 \quad (26c)$$

Based on the assumption of substitutability between routine labor and capital, i.e. ( $\gamma_r < \theta$ ), these comparative statics show that how the rise of the capital stock changes wage ratios and therefore the thresholds of educational frictions.

For example, the threshold of educational frictions with respect to the choice between routine and manual employment, i.e. secondary or primary education, is defined in equation (11) as:

$$\hat{\tau}_{ijt}^{mr} = x_{mjt} \left[ 1 - \frac{w_{mjt+1}}{w_{rjt+1}} \right],$$

where an increase in the wage ratio between manual and routine occupations (induced by falling capital prices) raises the threshold level of educational frictions for which an individual chooses secondary over primary education. Hence, less cross-generational transfers out of manual into routine occupation occur. While transitions from manual into routine occur at a lower rate, this wage ratio also impacts children whose parents work in routine occupations, where an increase in the wage ratio between manual and routine occupations leads to more transitions out of routine into manual employment.

Just the same pattern emerges when considering individuals choosing between secondary and tertiary education, i.e. routine and abstract employment.

$$\hat{\tau}_{ijt}^{ra} = \frac{x_{rjt} \left[ 1 - \frac{w_{rjt+1}}{w_{ajt+1}} \right]}{z - \frac{w_{rjt+1}}{w_{ajt+1}}},$$

$$\hat{\tau}_{ijt}^{ar} = \frac{x_{ajt} \left[ 1 - \frac{w_{rjt+1}}{w_{ajt+1}} \right]}{z - \frac{w_{rjt+1}}{w_{ajt+1}}}.$$



## A.2 Labor Demand

Labor demand for manual and abstract tasks is derived from equation (10 and plugging in wages based on the firm maximization problem, i.e. equations (5a) and (5c). As the argumentation is analogous, I will only derive labor demand for manual tasks in detail.

It holds that:

$$L_{m1t} = \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_{m-1}}} L_{m2t}. \quad (27)$$

Further, I use the labor market clearing condition ( $L_{it} = L_{i1t} + L_{i2t} \forall i$ ), where  $i \in (m, r, a)$ . Substituting the labor market clearing condition in for  $L_{m1t}$  and  $L_{m2t}$  separately, yields the following labor demands:

$$L_{m1t} = \frac{\left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_{m-1}}} L_{mt}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_{m-1}}}} \quad \text{and} \quad L_{m2t} = \frac{L_{mt}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_{m-1}}}} \quad (28)$$

Analogously, for abstract labor demand:

$$L_{a1t} = \frac{\left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_{m-1}}} L_{at}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_{m-1}}}} \quad \text{and} \quad L_{a2t} = \frac{L_{at}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_{m-1}}}} \quad (29)$$

Labor demand for routine labor in both cities is based on equation (5d). Manipulating for each city yields:

$$L_{r1t} = \left\{ \frac{1}{A_r} \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right] \right\}^{\frac{1}{\theta}} K_{1t} \quad (30)$$

$$L_{r2t} = \left\{ \frac{1}{A_r} \left[ \left( \frac{p_{kt}}{A_2 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{2t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right] \right\}^{\frac{1}{\theta}} K_{2t} \quad (31)$$

It is important to note that labor demand for manual and abstract depends on the ratios of exogenous productivity levels and the price of housing, whereas demand for routine tasks depends on local exogenous productivity and local capital demand.

### A.3 Productivity and Housing Prices

This sections proves that housing prices are higher in the more productive region. Starting from the housing equilibrium, where total housing supply  $H$  is constant for all locations  $j$ :

$$\sum_i h_{ijt} L_{ijt} = H \quad \forall j. \quad (32)$$

Substituting the equilibrium allocation for housing ( $h_{ijt+1}^* = \frac{\beta w_{ijt+1}}{p_{jt+1}}$ ) yields:

$$\sum_i w_{ijt} L_{ijt} = H \frac{p_{jt}}{\beta} \quad \forall j. \quad (33)$$

Combining this result for both regions, I obtain:

$$\frac{p_{1t}}{p_{2t}} = \frac{w_{m1t} L_{m1t} + w_{r1t} L_{r1t} + w_{a1t} L_{a1t}}{w_{m2t} L_{m2t} + w_{r2t} L_{r2t} + w_{a2t} L_{a2t}}. \quad (34)$$

Plugging in wages from equations (5a) to (5c) and rearranging gives:

$$\begin{aligned} (A_r L_{r1t}^\theta + A_k K_{1t}^\theta)^{\frac{\gamma_r - \theta}{\theta}} A_r L_{r1t}^\theta - \frac{p_{1t}}{p_{2t}} \frac{A_2}{A_1} (A_r L_{r1t}^\theta + A_k K_{2t}^\theta)^{\frac{\gamma_r - \theta}{\theta}} A_r L_{r2t}^\theta &= \\ = \left[ \frac{p_{1t}}{p_{2t}} \frac{A_2}{A_1} \right] \left[ A_m L_{m2t}^{\gamma_m} + A_a L_{a2t}^{\gamma_a} \right] - A_m L_{m1t}^{\gamma_m} - A_a L_{a1t}^{\gamma_a} & \quad (35) \end{aligned}$$

Plugging in the labor demands for manual and abstract in both regions from Appendix A.2, and manipulating yields:

$$\begin{aligned} (A_r L_{r1t}^\theta + A_k K_{1t}^\theta)^{\frac{\gamma_r - \theta}{\theta}} A_r L_{r1t}^\theta - \frac{p_{1t}}{p_{2t}} \frac{A_2}{A_1} (A_r L_{r1t}^\theta + A_k K_{2t}^\theta)^{\frac{\gamma_r - \theta}{\theta}} A_r L_{r2t}^\theta &= \\ = A_m \left( \frac{L_{m1t}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_m - 1}}} \right)^{\gamma_m} \left[ \frac{A_2}{A_1} \frac{p_{1t}}{p_{2t}} - \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{\gamma_m}{\gamma_m - 1}} &+ \\ + A_a \left( \frac{L_{a1t}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{1}{\gamma_a - 1}}} \right)^{\gamma_a} \left[ \frac{A_2}{A_1} \frac{p_{1t}}{p_{2t}} - \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \right]^{\frac{\gamma_a}{\gamma_a - 1}} & \quad (36) \end{aligned}$$

Now I focus on the LHS of equation (36) by starting to reconsider the wages for routine in both regions. Combining equation (5b) for both regions and the real wage equalization equation (10), I can show that:

$$(A_r L_{r2t}^\theta + A_k K_{2t}^\theta)^{\frac{\gamma_r - \theta}{\theta}} = \left( \frac{p_{2t}}{p_{1t}} \right)^\beta \frac{A_1}{A_2} (A_r L_{r1t}^\theta + A_k K_{1t}^\theta)^{\frac{\gamma_r - \theta}{\theta}} \left( \frac{L_{r1t}}{L_{r2t}} \right)^{\theta - 1} \quad (37)$$

Plugging this result into the LHS of equation (36) and some rearranging yields:

$$\underbrace{(A_r L_{r1t}^\theta + A_k K_{1t}^\theta)^{\frac{\gamma_r - \theta}{\theta}}}_{\text{A}} \underbrace{A_r L_{r1t}^\theta}_{\text{B}} \left[ 1 - \left( \frac{p_{1t}}{p_{2t}} \right)^{1 - \beta} \underbrace{\frac{L_{r2t}}{L_{r1t}}}_{\text{C}} \right] \quad (38)$$

Now we can substitute multiple terms from equation (38), which is equal to the LHS of equation (36) from other parts. Specifically, I concentrate on the first term ( $A$ ), the second term ( $B$ ) and the last term in the squared bracket ( $C$ ). For  $A$ , I can use (5d):

$$(A_r L_{r1t}^\theta + A_k K_{1t}^\theta)^{\frac{\gamma_r - \theta}{\theta}} = \frac{P_{kt}}{A_1 \gamma_r A_k} K_{1t}^{1-\theta} \quad (39)$$

Second, for  $B$ , from the labor demand for routine labor in region 1, I get:

$$A_r L_{r1t}^\theta = \left( \frac{P_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k K_{1t}^\theta \quad (40)$$

Third, for  $C$ , the derivation is somewhat longer. Starting with the acknowledging that the price for capital is the same for all regions, and therefore using equation (10), it holds that:

$$\frac{\frac{w_{r1t}}{P_{kt}}}{\frac{w_{r2t}}{P_{kt}}} = \left( \frac{P_{1t}}{P_{2t}} \right)^\beta \quad (41)$$

Plugging in the equations (5b) and (5d) for both regions and rearranging yields:

$$\frac{L_{r1t}}{L_{r2t}} = \left( \frac{P_{1t}}{P_{2t}} \right)^{\frac{\beta}{\theta-1}} \frac{K_{1t}}{K_{2t}} \quad (42)$$

Plugging in labor demands for routine tasks for both locations, i.e. equations (30) and (31), and rearranging

$$\left( \frac{P_{1t}}{P_{2t}} \right)^{\frac{\beta}{\theta-1}} = \frac{\frac{P_{kt}}{A_1 \gamma_r A_k} \frac{\theta}{\gamma_r - \theta} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k}{\frac{P_{kt}}{A_2 \gamma_r A_k} \frac{\theta}{\gamma_r - \theta} K_{2t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k} \quad (43)$$

Solving for  $K_{2t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}}$  yields:

$$K_{2t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} = \left( \frac{A_2}{A_1} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{P_{1t}}{P_{2t}} \right)^{\frac{\beta\theta}{\theta-1}} + \left[ 1 - \left( \frac{P_{1t}}{P_{2t}} \right)^{\frac{\beta\theta}{\theta-1}} \right] A_k \left( \frac{P_{kt}}{A_2 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} \quad (44)$$

I am also exploiting the labor market clearing condition for routine labor ( $L_{rt} = L_{r1t} + L_{r2t}$ ), where I also plug in the respective labor demands given in equations. Solving this for  $K_{2t}$ , yields:

$$K_{2t} = \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{P_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left[ \left( \frac{P_{kt}}{A_2 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{2t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}}} \quad (45)$$

Plugging equation (44) into the denominator of equation (45) and manipulation gives:

$$K_{2t} = \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{P_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left( \frac{P_{1t}}{P_{2t}} \right)^{\frac{\beta}{\theta-1}} \left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{P_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}}} \quad (46)$$

Plugging equation (46) into equation (42) and replacing  $L_{r2t} = L_{rt} - L_{r1t}$  gives:

$$\frac{L_{rt} - L_{r1t}}{L_{r1t}} = \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}. \quad (47)$$

Now that I have found expressions for the terms  $A$ ,  $B$  and  $C$  in equation (38) shown in equations (39), (40) and (47), respectively, I can plug them into equation (36):

$$\begin{aligned} & \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\gamma_r}{\gamma_r - \theta}} K_{1t}^{\frac{\gamma_r(1-\theta)}{\gamma_r - \theta}} - \frac{p_{kt}}{A_1 \gamma_r} K_{1t} \right] \times \\ & \times \left\{ \frac{\left( 1 + \left( \frac{p_{1t}}{p_{2t}} \right)^{1-\beta} \right) \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t} - \left( \frac{p_{1t}}{p_{2t}} \right)^{1-\beta} L_{rt} A_r^{\frac{1}{\theta}}}{\left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}} \right\} = \\ & = A_m \left( \frac{L_{mt}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^{\beta} \right]^{\frac{1}{\gamma_m - 1}}} \right)^{\gamma_m} \left[ \frac{A_2 p_{1t}}{A_1 p_{2t}} - \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^{\beta} \right]^{\frac{\gamma_m}{\gamma_m - 1}} + \\ & + A_a \left( \frac{L_{at}}{1 + \left[ \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^{\beta} \right]^{\frac{1}{\gamma_a - 1}}} \right)^{\gamma_a} \left[ \frac{A_2 p_{1t}}{A_1 p_{2t}} - \frac{A_2}{A_1} \left( \frac{p_{1t}}{p_{2t}} \right)^{\beta} \right]^{\frac{\gamma_a}{\gamma_a - 1}} \end{aligned} \quad (48)$$

Equation (43) is essentially equivalent to Eeckhout et al. (2019, Eq. F.1).

I can now commence the proof for housing prices dependent on productivity advantages in location 1. I am working with contradictions. Assume  $A_2 > A_1$  and  $p_{1t} > p_{2t}$ . This implies that the RHS of equation (48) is greater than zero as  $\beta \in (0, 1)$ , so the LHS also has to be greater than zero. The first term in squared brackets is also greater than zero - see equations (39) and (40) -, so I have to investigate the term in curly brackets in equation (48): For the inequality to hold, it has to be greater than zero, so after some rearrangement, I obtain:

$$K_{1t} \left( \frac{p_{2t}}{p_{1t}} \right)^{\beta - 1} > \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left[ \left( \frac{p_{kt}}{A_2 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}} \quad (49)$$

Dividing the last equation by  $\frac{p_{1t}}{p_{2t}}^{\frac{\beta}{1-\theta}}$ , this yields:

$$K_{1t} \left( \frac{p_{2t}}{p_{1t}} \right)^{1 + \frac{\beta\theta}{1-\theta}} > \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\frac{p_{1t}}{p_{2t}}^{\frac{\beta}{1-\theta}} \left[ \left( \frac{p_{kt}}{A_2 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}} \quad (50)$$

When combining equations (44) and (46), I obtain:

$$\begin{aligned}
& \left\{ \frac{L_{rt}A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1\gamma_r A_k} \right)^{\frac{\theta}{\gamma_r-\theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left( \frac{p_{1t}}{p_{2t}} \right)^{\frac{\beta}{1-\theta}} \left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} \left( \frac{p_{kt}}{A_1\gamma_r A_k} \right)^{\frac{\theta}{\gamma_r-\theta}} - A_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} = \\
& = \left( \frac{A_2}{A_1} \right)^{\frac{\theta}{\gamma_r-\theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} \left( \frac{p_{1t}}{p_{2t}} \right)^{\frac{\beta\theta}{1-\theta}} + \left[ 1 - \left( \frac{p_{1t}}{p_{2t}} \right)^{\frac{\beta\theta}{1-\theta}} \right] A_k \left( \frac{p_{kt}}{A_2\gamma_r A_k} \right)^{\frac{\theta}{\theta-\gamma_r}}. \quad (51)
\end{aligned}$$

This equation is essentially equivalent to Eeckhout et al. (2019, Eq. F.2). I can infer from equation (51) that, given the assumptions about exogenous productivity differences and prices, i.e.  $A_2 > A_1$  and  $p_{1t} > p_{2t}$ , that the last term is smaller than zero. Consequently it has to hold:

$$K_{1t} \left( \frac{p_{2t}}{p_{1t}} \right)^{\frac{\beta\theta}{1-\theta} \frac{\theta-\gamma_r}{(1-\gamma_r)\theta}} < \frac{L_{rt}A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1\gamma_r A_k} \right)^{\frac{\theta}{\gamma_r-\theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left( \frac{p_{1t}}{p_{2t}} \right)^{\frac{\beta}{1-\theta}} \left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} \left( \frac{p_{kt}}{A_1\gamma_r A_k} \right)^{\frac{\theta}{\gamma_r-\theta}} - A_k \right]^{\frac{1}{\theta}}} \quad (52)$$

Equations (50) and (52) allow me to focus on the exponents of the LHS, respectively. For the assumptions set out at the beginning of the proof (i.e.  $A_2 > A_1$  and  $p_{1t} > p_{2t}$ ) to hold, I need that both inequalities hold, indicating that the exponent of equation (52) is smaller than that of equation (50) given that  $\frac{p_{2t}}{p_{1t}} < 1$ . Thus, by subtracting the latter from the former, it should give a negative value:

$$1 + \frac{\beta\theta}{1-\theta} - \frac{\beta\theta}{1-\theta} \frac{\theta-\gamma_r}{(1-\gamma_r)\theta} = 1 + \frac{\beta\theta}{1-\theta} \left( 1 - \frac{\theta-\gamma_r}{(1-\gamma_r)\theta} \right) = 1 + \frac{\beta\gamma_r}{1-\gamma_r} > 0 \quad (53)$$

Hence, I get a contradiction from equations 50 and 52, and it holds that the more productive region has higher housing prices. In formal terms, this proves that  $p_{1t} > p_{2t}$  if  $A_1 > A_2$ .

I further show that it holds: If  $A_1 > A_2$ , then  $\frac{A_1}{A_2} \left( \frac{p_{2t}}{p_{1t}} \right)^{\beta} > 1$ .

Starting again with equation (48) and the squared bracket, and a proof by contradiction, I assume that  $A_1 > A_2$  and  $\frac{A_1}{A_2} \left( \frac{p_{2t}}{p_{1t}} \right)^{\beta} < 1$ . This assumption implies that the squared bracket is positive, rendering its RHS also greater than 0:

$$\frac{A_2 p_{1t}}{A_1 p_{2t}} - \left[ \frac{A_1}{A_2} \left( \frac{p_{1t}}{p_{2t}} \right)^{\beta} \right]^{\frac{\gamma_i}{\gamma_i-1}} > 0, \quad \text{with } i \in \{m, r\}. \quad (54)$$

This implies that the LHS of equation (48) must also be positive. Following the previous argumentation, the inequality of equation (50) also has to hold. Also, in equation (51), we know that the last term is negative as  $p_{1t} > p_{2t}$ . Therefore: Using the RHS of both these equations, and dividing yields:

$$\frac{L_{rt}A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1\gamma_r A_k} \right)^{\frac{\theta}{\gamma_r-\theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left( \frac{p_{1t}}{p_{2t}} \right)^{\frac{\beta}{1-\theta}} \left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r-\theta}} \left( \frac{p_{kt}}{A_1\gamma_r A_k} \right)^{\frac{\theta}{\gamma_r-\theta}} - A_k \right]^{\frac{1}{\theta}}} > \left( \frac{A_2}{A_1} \right)^{\frac{1}{1-\gamma_r}} K_{1t} \left( \frac{p_{1t}}{p_{2t}} \right)^{\frac{\beta}{1-\theta} \frac{\gamma_r-\theta}{1-\gamma_r}} \quad (55)$$

Using the last inequality and dividing it by the last inequality gives:

$$\frac{\text{RHS}(50)}{\text{RHS}(55)} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_r}} \left(\frac{p_{2t}}{p_{1t}}\right)^{1+\frac{\beta\theta}{(1-\theta)}} \left[1 - \frac{\gamma_r - \theta}{\theta(1-\gamma_r)}\right] \quad (56)$$

After some rearranging, I can write the previous equation as:

$$\frac{\text{RHS}(50)}{\text{RHS}(55)} = \left[\frac{A_1}{A_2} \left(\frac{p_{2t}}{p_{1t}}\right)^{1-\gamma_r(1-\beta)}\right]^{\frac{1}{1-\gamma_r}}. \quad (57)$$

Analogous to the exercise above of comparing exponents, I can show that:  $1 - \gamma_r(1 - \beta) > \beta$ . As  $p_{1t} > p_{2t}$ , then it holds that:

$$\frac{A_1}{A_2} \left(\frac{p_{2t}}{p_{1t}}\right)^{1-\gamma_r(1-\beta)} < \frac{A_1}{A_2} \left(\frac{p_{2t}}{p_{1t}}\right)^{\beta} < 1, \quad (58)$$

where the last inequality comes from the initial assumption to work towards contradiction. As  $\gamma_r \in (0, 1)$ , then  $\frac{1}{1-\gamma_r} > 0$ . As a consequence,  $\frac{\text{RHS}(50)}{\text{RHS}(55)} < 0$ . This means that both inequalities cannot be satisfied and therefore I have a contradiction.

Formally, this means:

$$\frac{A_1}{A_2} \left(\frac{p_{2t}}{p_{1t}}\right)^{\beta} > 1. \quad (59)$$

#### A.4 Productivity and Population Size

The calculation is the same for both manual and abstract labor, I will show this with the example of manual labor. Labor demand for manual tasks in location 1 is given by equation (28):

$$L_{m1t} = \frac{\left[ \frac{A_1}{A_2} \left( \frac{p_{2t}}{p_{1t}} \right)^\beta \right]^{\frac{1}{1-\gamma_m}} L_{mt}}{1 + \left[ \frac{A_1}{A_2} \left( \frac{p_{2t}}{p_{1t}} \right)^\beta \right]^{\frac{1}{1-\gamma_m}}} \quad (60)$$

As shown in Appendix A.3, it holds that  $\frac{A_1}{A_2} > \left( \frac{p_{2t}}{p_{1t}} \right)^\beta$ . Using these results for manual labor demand in location 1 yields:

$$L_{m1t} > \frac{L_{mt}}{2}, \quad (61)$$

as, by assumption,  $\frac{1}{1-\gamma_m} > 1$ . This implies that  $L_{m1t} > L_{m2t}$  and, as stated above analogously for abstract labor,  $L_{a1t} > L_{a2t}$ .

Now, location 1 can only be smaller if  $L_{r2t} > L_{r1t}$ . To show that this is not the case, I will be working with a contradiction.

Re-consider equation (51). I have shown that  $p_{1t} > p_{2t}$  if  $A_1 > A_2$ . This means that the second term of this equation is negative, and it holds:

$$\left\{ \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} > \left[ \left( \frac{p_{1t}}{p_{2t}} \right)^\beta \frac{A_2}{A_1} \right]^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \quad (62)$$

Going back to equation (47), and using the assumption  $L_{r2t} > L_{r1t}$  (implying  $\frac{L_{rt} - L_{r1t}}{L_{r1t}} > 1$ ), it also holds:

$$K_{1t} > \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}}} \quad (63)$$

Manipulation of the previous equation implies:

$$K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} > \left\{ \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \quad (64)$$

Combining equations (62) and (64) implies that:

$$K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} > \left\{ \frac{L_{rt} A_r^{\frac{1}{\theta}} - \left[ \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}} K_{1t}}{\left[ K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \left( \frac{p_{kt}}{A_1 \gamma_r A_k} \right)^{\frac{\theta}{\gamma_r - \theta}} - A_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} > \left[ \left( \frac{p_{2t}}{p_{1t}} \right)^\beta \frac{A_1}{A_2} \right]^{\frac{\theta}{\gamma_r - \theta}} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \quad (65)$$

As shown above, it holds that  $\frac{A_1}{A_2} \left( \frac{p_{2t}}{p_{1t}} \right)^\beta > 1$ . This implies that, as the last term the first term of the third component in equation (65) is greater than one as  $\gamma_r < \theta$ . Hence, I found a contradic-

tion, and it must hold that  $L_{r1t} > L_{r2t}$ .

Ultimately, it holds that the share of each labor type in location 1 is higher than in location 2, hence it holds that the size of location 1 is greater than that of location 2, i.e.  $S_1 > S_2$  as  $S_j = \sum_i L_{ijt} \forall j$ .



## A.5 Productivity and Demand for Capital

Following the proof from Appendix A.3, i.e. prices are higher in region 1 compared to region 2 ( $p_{1t} > p_{2t}$ ) if  $A_1 > A_2$ . The price for capital is the same for all regions, therefore, using equation (10), it holds: :

$$\frac{\frac{w_{r1t}}{p_{kt}}}{\frac{w_{r2t}}{p_{kt}}} = \left(\frac{p_{1t}}{p_{2t}}\right)^\beta \quad (66)$$

Plugging in the equations (5b) and (5d) for both regions and rearranging yields:

$$\frac{L_{r1t}}{L_{r2t}} = \left(\frac{p_{1t}}{p_{2t}}\right)^{\frac{\beta}{\theta-1}} \frac{K_{1t}}{K_{2t}} \quad (67)$$

Plugging in labor demands for routine tasks for both locations, i.e. equations (30) and (31), and rearranging

$$\left(\frac{p_{1t}}{p_{2t}}\right)^{\frac{\beta}{\theta-1}} = \frac{\frac{p_{kt}}{A_1 \gamma_r A_k} \frac{\theta}{\gamma_r - \theta} K_{1t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k}{\frac{p_{kt}}{A_2 \gamma_r A_k} \frac{\theta}{\gamma_r - \theta} K_{2t}^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} - A_k} \quad (68)$$

As shown above, it holds that  $p_{1t} > p_{2t}$  if  $A_1 > A_2$  and by assumption,  $\gamma_r < 1$ , then the LHS of equation (68) is smaller than 1. Hence, the RHS also has to be smaller than 1. Setting the LHS smaller than zero, and simplifying then yields:

$$\left(\frac{A_1}{A_2}\right)^{\frac{\theta}{\gamma_r - \theta}} > \left(\frac{K_{1t}}{K_{2t}}\right)^{\frac{(1-\gamma_r)\theta}{\gamma_r - \theta}} \quad (69)$$

Remember that  $\gamma_r < \theta$ , hence the inequality sign turns around when cancelling out exponents as far as possible. Then it holds that:

$$\frac{A_1}{A_2} < \left(\frac{K_{1t}}{K_{2t}}\right)^{1-\gamma_r} \quad (70)$$

By assumption, if  $A_1 > A_2$ , the LHS of equation (70) is greater than 1. Subsequently, the RHS must also be greater than 1, and as  $\gamma_r < 1$ , the RHS is only greater than one if  $K_{1t} > K_{2t}$ . Consequently, capital demand is higher in the location with higher exogenous productivity.

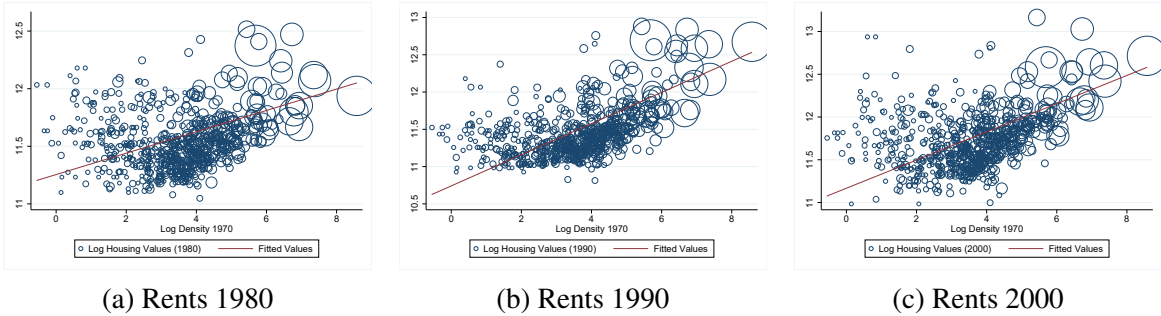
# Appendix B - Empirical Evidence I

Table B1: Correlations of Density and Human Capital with ICT Investment per Worker

	(1)	(2)	(3)
Log Population (1970)	0.57*** (19.64)	1.08*** (26.11)	2.21*** (25.22)
Constant	4.38*** (26.89)	3.06*** (13.15)	6.23*** (12.67)
Obs.	736	736	736
R <sup>2</sup>	0.34	0.48	0.46

Notes: The dependent variable is investment in computer capital per worker in USD. Column (1) show ICT investment in 1982, column (2) for 1992, and column (3) for 1997. Columns (1) to (3) regress ICT investment per worker on log density in 1970. The coefficients are equal to the slope of the fitted line in Figure 7. Observations are weighted by population share in 1970. t-statistics are shown in brackets. \* denotes 10% significance, \*\* denotes 5% significance, \*\*\* denotes 1% significance.

Figure B1: Density and Housing Values



The scatter plots show the correlations between historical levels of density (expressed by its log) and housing values in various years on the commuting-zone level. As expected from the model, the slope is positive. Observations are weighted by their 1980 population share.

Table B2: Correlations of Density and Educational Polarization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1980	1980	1990	1990	2000	2000	2010	2010	2018	2018
	Median = 6	Median = 7	Median = 6	Median = 7	Median = 6	Median = 7	Median = 6	Median = 7	Median = 6	Median = 7
Log Density (1970)	0.29*** (12.56)	0.27*** (10.00)	0.02 (0.49)	0.47*** (15.08)	0.42*** (13.64)	0.56*** (16.42)	0.07** (2.07)	0.48*** (19.34)	0.08** (2.32)	0.30*** (10.94)
Constant	-0.95*** (-8.56)	0.23 (1.46)	-0.26** (-2.13)	-1.06*** (-6.31)	-1.37*** (-10.57)	-1.34*** (-7.07)	-0.51*** (-4.16)	-1.36*** (-10.35)	-0.81*** (-6.60)	-0.62*** (-5.13)
Obs.	658	77	439	297	507	229	342	393	263	382
R <sup>2</sup>	0.19	0.57	0.00	0.44	0.27	0.54	0.01	0.49	0.02	0.24

Notes: The dependent variable is the Allison-Foster index of educational polarization. It is computed with Decennial Censuses on the commuting zone level. The table regresses the indices of all decades from 1980 to 2018 on log density in 1970. In all decades, the table distinguishes between commuting zones with median equal to 6 or 7, i.e. high school graduate or one year of college, respectively. Observations are weighted by population share in 1970. p-values are shown in brackets. \* denotes 10% significance, \*\* denotes 5% significance, \*\*\* denotes 1% significance.

Table B3: Parental Transfers to Children with six occupational groups (1988)

Occupational Group	(1) Parental Income	(2) Money from Parents (USD)	(3) Money from Parents (%)	(4) Time Help from Parents
Service	35936.69	217.96	.49	178.33
Transport/Construct/Mechanic	54574.99	264.13	.49	133.88
Cleric/Retail	55876.16	509.36	.95	78.64
Operators	54598.2	586.7	.93	137.07
Production/Craft	69030.2	436.3	.62	61.37
Managers/Professionals	79792.84	600	.79	73.01

*Notes:* This table shows parental income and transfers of time and money to their children. It shows that income is rising with occupational group, and so are monetary transfers by parents, both in absolute and percent. Help in terms of hours declines with occupational groups. The data is taken from the in 1988 by the PSID.

Table B4: Parental Transfers to Children with six occupational groups (2013)

Occupational Group	(1) Parental Income	(2) Money from Parents (USD)	(3) Money from Parents (%)	(4) Support School (%)	(5) Time Help from Parents
Service	39757.98	78.59	.2	12.56	58.6
Transport/Construct/Mechanic	62125.93	613.63	.89	16.2	78.36
Cleric/Retail	67713.45	659.89	1.08	23.87	84.53
Operators	55508.43	389.71	.71	13.84	79.95
Production/Craft	69596.14	298.37	.39	21.35	74.5
Managers/Professionals	90523.03	1036.99	1.07	34.56	87

*Notes:* This table shows parental income and transfers of time and money to their children. It shows that income is rising with occupational group, and so are monetary transfers by parents, both in absolute and percent. Help in terms of hours declines with occupational groups. The data is taken from the in 2013 by the PSID.