Like Father, Like Son: Occupational Choice, Intergenerational Persistence and Misallocation

Salvatore Lo Bello* and Iacopo Morchio‡

October 2018

Abstract

We develop a dynamic quantitative model of occupational choice and search frictions with multiple channels of intergenerational transmission (comparative advantage, social contacts and preferences), and use it to decompose the occupational persistence observed in the UK. In the model, workers who choose their father's occupation find jobs faster and earn lower wages, which is consistent with patterns found in UK data. Quantitatively, parental networks account for 78% of total persistence. Shutting down parental networks or the transmission of preferences improves the allocation of workers and thus yields welfare gains, while removing the transmission of comparative advantage generates welfare losses.

Keywords: Comparative Advantage, Labor Productivity, Mismatch, Occupational Mobility, Social Contacts.
JEL Classification Numbers: J24, J62, J64.

*Bank of Italy. E-mail: salvatore.lobello@esterni.bancaditalia.it

†University of Vienna. E-mail: iacopo.morchio@univie.ac.at

‡We thank Andrès Erosa for his guidance and support. We are also grateful to Arpad Abraham, Esther Ann Boler, George Borjas, Carlos Carrillo–Tudela, Betsy Caucutt, Klaus Desmet, Juanjo Dolado, Javier Fernandez–Blanco, Manolis Galenianos, Manuel Garcia–Santana, Pedro Gomes, Belen Jerez, Leo Kaas, Tim Kehoe, Matthias Kredler, Rody Manuelli, Claudio Michelacci, Kurt Mitman, Alessio Moro, Josè-Vicente Rodriguez Mora, Rachel Ngai, Juan Pablo Nicolini, Theodore Papageorgiou, Christopher Rauh, Mark Rosenzweig, Raoul Santaeulalia–Lloepis, Itay Saporta, Stefanie Stantcheva, Jan Stuhler, Giorgio Topa, Ludo Visschers, Felix Wellschmied, Minchul Yum and Emircan Yurdagul, as well as Annalisa Loviglio for a discussion of an earlier version of this paper, and seminar and conference participants at the PhD Student UC3M Workshop, Macro Breakfast (University of Vienna), ENTER Jamboree 2016 (Madrid), Intergenerational Mobility Workshop (Madrid), VII IIBEO Workshop 2016 (Alghero), XXI Workshop in Dynamic Macro 2016 (Vigo), Doctoral Workshop on Quantitative Dynamic Economics 2016 (Marseille), LACEA 2016 (Bogotá), Universidad del Rosario, World Bank, University of Edinburgh, I UC3M-CEMFI Macro PhD Workshop, CSEF (Naples), SMU (Dallas), University of Montreal, York University, University of Bologna, Bank of Italy, University of Surrey, HSE (Moscow), New Economic School (Moscow), 7th European SaM Conference, University of Kent and others for useful comments and suggestions. Financial support from the Spanish Ministry of Education is gratefully acknowledged.

1 Introduction

It is well known that a number of economic outcomes are correlated across generations, most notably income, education and occupational choice. Such persistence is commonly believed to represent a failure of the equality of opportunity principle, besides being potentially symptomatic of an underlying misallocation of resources and talents (Mora 2009; Caselli & Gennaioli 2013; Güell *et al.* 2015). In particular, a large degree of persistence in occupational choice may reflect the presence of barriers of various types in the labour market, implying a suboptimal allocation of workers to jobs. However, a quantitative theory of occupational persistence, that would help us understand whether or not persistence is indeed associated with inefficiencies, has not been developed yet.

In this paper, we study how intergenerational occupational persistence and labor misallocation are related. We demonstrate that when persistence stems from a number of sources, it is crucial to measure the quantitative importance of each and how they interact with one another. To do so, we develop a dynamic model of occupational choice and search frictions that features multiple channels of intergenerational transmission, and use it to decompose the occupational persistence observed in UK data. Our model provides novel insights on the impact of different sources of persistence on the sorting of workers in the labor market, and therefore the aggregate level of efficiency and welfare.

To illustrate our main mechanism, we first develop a simple two–period model, in which both abilities and social contacts are transmitted across generations. We show that persistence can indeed be a sign of misallocation if parents help their offspring find a job faster in their current occupation, which is not necessarily where their offspring's comparative advantage lies. In our model, workers optimally choose their occupation, so that mismatch is not always detrimental to welfare. However, we find the equilibrium to generally be inefficient, as workers do not internalize that: i) search frictions interact with the level of mismatch present in the economy, and firms offer fewer jobs if they expect workers to be more mismatched and consequently less productive;[1] and ii) mismatch has dynamic effects, via intergenerational transmission. We also analytically establish that higher levels of persistence can be associated with either higher or lower welfare, depending on whether they are the result of stronger transmission of productivity or social contacts across generations.

To assess the quantitative importance of the different channels of persistence, we then extend our theory to a dynamic model that embeds a third channel of transmission (i.e. preferences)[2] and allows for mobility over the life–cycle. We investigate the quantitative

[1]This channel is present also in Bentolila *et al.* (2010).

[2]Studies on intergenerational transmission of preferences and work attitudes include Dohmen *et al.* (2011), Paola (2013) and Bisin & Verdier (2005). Research on how these transmitted traits affect occupational choice

importance of the three different forces in generating occupational persistence, and find that parental networks can account for about 78% of persistence, whereas transmission of comparative advantage and preferences account for 19% and 10%, respectively.[3] The very large role played by parental networks depends on the fact that they strongly interact with the other two factors generating persistence, thus acting as a multiplier. Moreover, we demonstrate that a model in which only productive abilities are transmitted across generations (which we call the *restricted model*) falls short in accounting for several key pieces of evidence and in general provides a much worse fit to the data.

The theory delivers two main predictions for agents who choose to pursue the same occupation as their parents (*occupational followers*) relative to those who do not (*occupational movers*). First, the former find jobs at a faster rate. Second, they earn lower wages. We find confirmation of both implications using data from the British Household Panel Survey (1991–2008) on male workers and their fathers. This dataset allows us to observe labor market transitions, occupational affiliation and wages of both fathers and sons, together with a large number of other covariates.

We estimate that the degree of occupational persistence in the UK is substantial: at the 1–digit level, a worker is 72% more likely to be employed in a given occupation if his father is also currently employed in that occupation. When we test the model's predictions, we find that occupational followers exhibit monthly job–finding rates that are 5–6 p.p. higher than those of occupational movers. Given that the average in–sample job–finding rate is about 12.5 p.p., being in the same occupation as one's father yields a substantial advantage in terms of unemployment risk (robust to the inclusion of individual fixed effects). Regarding wages, we find that occupational followers exhibit large discounts (between 7 and 14 log points) relative to occupational movers. These results are shown to be robust to alternative definitions of the *occupational followers* group, using information on the entire labor market careers as well as only contemporaneous information. Finally, we also document that sons of high–wage fathers are more likely to be occupational followers, a fact that supports our theory, which is based on comparative advantage (rather than absolute advantage).

We calibrate the quantitative model in order to match the above–mentioned key pieces of empirical evidence, along with several other moments of the UK economy. In particular, the parameters governing intergenerational transmission are pinned down as follows: the parental networks transmission replicates the job–finding rate premium of followers, the comparative advantages replicates the differential in persistence by parental wage, and the preferences

includes Doepke & Zilibotti (2008), Caner & Okten (2010) and Escriche (2007).

[3]The sum of the effects of the three transmission channels exceeds 100% because they are endogenously correlated.

transmission is the residual channel of occupational persistence. We use the quantitative model to assess how each source of persistence affects welfare. When we shut down parental networks or the transmission of preferences, welfare improves. This is due to the improved (i.e. more aligned with productive advantages) allocation of workers to occupations: output per worker goes up, and firms react by posting more vacancies per worker than previously. In contrast, when we shut down the transmission of abilities, the reduction in persistence is accompanied by a reduction in welfare, driven by the inferior allocation of the workforce, which in turn also leads to less firm entry. However, the changes in welfare are relatively small, with a magnitude of 0.1% in consumption equivalent variation (CEV).

We also investigate the role of search frictions and find that the impact of parental networks on the allocation becomes negligible as frictions tend to zero; also, an increase in the severity of frictions in our model simultaneously generates a rise in the unemployment rate and a drop in labor productivity. Finally, in our model more generous unemployment benefits imply that workers are less likely to choose the same occupation as their father in order to reduce the probability of unemployment. This implies that, in addition to reducing occupational persistence, increasing benefits can potentially be welfare–improving, since the allocation of the workforce may also improve. Nonetheless, we find that an increase in unemployment benefits triggers a reallocation towards preferences rather than comparative advantages. Therefore, the overall effect on welfare is slightly negative. For instance, an increase in unemployment benefits of 25% yields a decrease in welfare of 0.3% (CEV). Importantly, the restricted model yields very different implications for this policy change by underestimating the negative welfare effects of increased benefits.

While there is a significant amount of research on income persistence across generations,[4] work on occupational choice is far scarcer in the literature.[5] We contribute to this literature along several dimensions: First, we add to the empirical literature on occupational persistence across generations (Constant & Zimmermann 2004, Hellerstein & Morrill 2011, Ermisch & Francesconi 2002, Di Pietro & Urwin 2003, Long & Ferrie 2013) by documenting new facts on labor market outcomes of occupational followers. We show that, relative to other observationally equivalent workers, they find jobs faster but earn lower wages. Moreover, we

[4]The first important contribution in the intergenerational literature was Becker & Tomes (1979); more recently, Solon (1992), Solon (2002), Corak (2006), Hertz (2006) and Björklund & Jäntti (2009) have documented persistence in income. For persistence in education, see Chevalier *et al.* (2009); and for occupational persistence, see Hout & Beller (2006), Constant & Zimmermann (2004), Escriche (2007), Eberharter (2008) and Dustmann (2004). Two reviews of the literature can be found in Black & Devereux (2010) and Ermisch *et al.* (2012).

[5]The economic literature has primarily focused on the study of income, while the sociological literature, pioneered by Blau & Duncan (1967), has focused on occupational persistence. More recent contributions include Stier & Grusky (1990), Checchi (1997) and Andres *et al.* (1999).

provide new estimates of the likelihood of belonging to the same occupational category as one's father using contemporaneous information based on monthly transitions.

Second, our study bridges the literature on the determinants of occupational choice (Miller 1984; McCall 1991; Keane & Wolpin 1997; Papageorgiou 2014; Carrillo-Tudela & Visschers 2014), its consequences for inequality (Kambourov & Manovskii 2009) and unemployment duration (Wiczer 2014) and the literature on occupational persistence and career following (Laband & Lentz 1983). Our paper is one of the few to adopt a theoretical and quantitative perspective on occupational choice across generations,[6] thus providing novel insights into how persistence maps to efficiency and aggregate welfare. We stress that in this paper we focus on the determinants of occupational persistence and how this affects aggregate outcomes, while we do not study the connection between occupational choice persistence and earnings persistence across generations.[7]

Third, we relate to the literature on social networks in the labor market (for instance, see Horváth 2014 and Galenianos 2014)[8] and particularly the transmission of contacts or related advantages across generations (Corak & Piraino 2011, Kramarz & Skans 2014, Pellizzari *et al.* 2011, Lentz & Laband 1989 and Aina & Nicoletti 2014). Fourth, our paper investigates the possibility of misallocation of the labor force due to socially suboptimal occupational choice, as in Bentolila *et al.* (2010), Hsieh *et al.* (2013) and Munshi & Rosenzweig (2016). In the intergenerational context, the occurrence of socially suboptimal occupational choice has been studied in Caselli & Gennaioli (2013) and Celik (2015). Finally, the decomposition exercise undertaken here is very close in spirit to those of income persistence across generations, such as Restuccia & Urrutia (2004), Lee & Seshadri (2014), Abbott *et al.* (2013) and Gayle *et al.* (2015).

The rest of the paper is organized as follows: Section 2 develops a two–period model of occupational choice and intergenerational transmission. Section 3 presents empirical evidence on occupational persistence. Section 4 outlines the dynamic quantitative model, which is calibrated and used for counterfactual experiments in Section 5. Finally, Section 6 concludes.

[6]A quantitative perspective is also adopted by Sinha (2014), who studies how borrowing constraints affect occupational choices and how this mechanism can be important in understanding persistence in developing countries.

[7]Doing so would require us to model persistence in general ability and to allow for heterogeneous occupations, while our model features homogeneous occupations. We leave this question to further research.

[8]Other contributions include Granovetter (1973), Montgomery (1991), Calvó-Armengol & Jackson (2007), Pellizzari (2010), Cingano & Rosolia (2012), Hensvik & Skans (2013), Topa (2001) and Dustmann *et al.* (forthcoming).

2 A Simple Two–Period Model

The model is based on Bentolila *et al.* (2010), which is a static version of the standard search model à la Diamond-Mortensen–Pissarides. We add the following features to their framework: i) overlapping generations that are connected to each other; ii) a notion of occupational persistence across generations; and iii) two different sources of persistence across generations.

The economy is populated by a measure 2 of risk–neutral individuals who live for two periods (a young and an old phase of life). The population has the typical OLG structure, with each young individual being connected to one old individual (who we call "his father"). There are two occupations denoted by $j \in \{A, B\}$, each of which corresponds to a separate labor market. The only choice made by individuals is in which occupation to search for a job.

In each of the two markets, matches between unemployed workers and vacancies take place according to a matching function $M(U_j, V_j)$, where U_j represents the total number of search efficiency units, and V_j is the number of vacancies in the market. We define the probability of finding a job for an unemployed worker exerting 1 unit of search efficiency as $p(\theta_j) = \frac{M(U_j,V_j)}{U_j}$, where $\theta_j = \frac{V_j}{U_j}$ represents the labor market tightness. Conversely, $q(\theta_j) = \frac{M(U_j,V_j)}{V_j}$ represents the job-filling rate (probability of a firm meeting a worker).

Upon matching, the generated surplus is split according to Generalized Nash Bargaining, where the worker's bargaining power is equal to β. If unmatched, workers receive an unemployment benefit (normalized to 0 without loss of generality). We assume that search takes place during the young phase, while production takes place only during the old phase.[9] As a result, equilibrium wages are equal to a share β of the worker's productivity.

Workers are heterogeneous in their occupation–specific productive advantage and in the endowment of social contacts. A worker without productive advantage produces y, whereas he produces $(1+a)y$ when he can exploit his productive advantage, where $a > 0$. We assume that there exist two types of workers, with $\tau = \{A, B\}$ denoting the occupation in which each type has his productive advantage.[10] Types are transmitted across generations according to the following transition probabilities: $P(\tau = A|\tau^F = A) = P(\tau = B|\tau^F = B) = \rho$ (we denote the father's variables with an F superscript), with $\rho \in [0, 1]$. Due to the symmetry of the transition matrix, the long–run distribution is characterized by $P(\tau = A) = P(\tau = B) = \frac{1}{2}$, so that types are independent if $\rho = \frac{1}{2}$. In what follows, we assume that $\rho > \frac{1}{2}$,

[9] These assumptions are made for simplicity, since they make the model static.

[10] We could allow the comparative advantage to vary, depending on the occupation chosen by the worker, and this would not alter the qualitative predictions of the model. The crucial assumption underlying our results is that ability types change more slowly than contacts.

that is an individual has a tendency to be of the same type as his father, on average. Importantly, in this framework there is no productive *absolute* advantage. This is consistent with Papageorgiou (2014) who argues that comparative advantage is the most important component in occupational choice.

Social contacts are occupation–specific and help individuals find jobs by increasing the efficiency units of search. We assume that young individuals have more social contacts in their father's occupation: one possible interpretation is that they can exploit their father's networks if they decide to enter the same occupation. Without loss of generality, we normalize the network of young workers who do not search in the same occupation as their father to zero. Thus, their efficiency units of search are equal to $1 + \tilde{n}$, where $\tilde{n}$ represents the size of the network they can exploit. We further assume that all workers, when passing from the young to the old phase, accumulate a network of size N with probability μ and of size n otherwise, with $\mu \in [0, 1]$. To sum up:

$$\tilde{n} = \begin{cases} 0 & \text{if } o \neq o^F \\ N & \text{if } o = o^F \qquad \text{w.p.} \qquad \mu \\ n & \text{if } o = o^F \qquad \text{w.p.} \quad (1-\mu) \end{cases}$$

where o denotes the occupation chosen, and o^F denotes the father's occupation. Finally, we also assume that $n < a < N$ (Assumption A), in order to induce different groups of workers to make different occupational choices.

2.1 Equilibrium characterization

We focus on steady state symmetric equilibria, given that the two markets have the same fundamentals.[11] In equilibrium, each market attracts a measure 1 of workers, with an identical composition of labor productivity and contacts endowment. In order to facilitate the exposition, we will sometimes frame the discussion using occupation A, given that the same results hold for occupation B.

There is a subset of workers in this economy for which the occupational choice is trivial. In other words, all those workers whose productive advantage lies in their father's occupation

[11] Thanks to the symmetry of the equilibrium, we will frequently use the following property: $P(A|B) = P(B|A)$ whenever A and B are events defined over the distribution of types and occupations (which have equal mass in equilibrium).

will simply choose that occupation. We calculate their mass as follows:[12]

$$\begin{aligned} P(\tau = A|o^F = A) = &P(\tau = A|o^F = A; \tau^F = A)P(\tau^F = A|o^F = A)+ \\ &P(\tau = A|o^F = A; \tau^F = B)P(\tau^F = B|o^F = A), \end{aligned} \tag{1}$$

where o^F denotes the father's occupation. Noting that type transmission does not depend on the occupational affiliation of the father, and after defining $P(o^F = A|\tau^F = A) = m$ (an endogenous object yet to be determined), we can simplify the previous expression to the following:

$$P(\tau = A|o^F = A) = \rho m + (1 - \rho)(1 - m) \equiv \Psi(m), \tag{2}$$

where we have used the fact that $P(\tau^F = A|o^F = A) = P(o^F = A|\tau^F = A)$. Here, $(1 - m)$ is a measure of the *productive mismatch* in the economy, i.e. how poor the sorting between workers and occupations is along the productivity dimension. If $m = 1$, the occupational choice is perfectly aligned with the type, that is, all workers of a given type enter the same occupation, which is where they are most productive. Moreover, $\Psi(m)$ denotes the mass of young workers who do not face any tradeoff in their occupational decision.

In order to solve for the equilibrium m, we first look at the optimal occupational choice of the workers who face the tradeoff. These are the workers whose productive advantage lies in one occupation but whose father (and his social contacts) is employed in the other. For instance, consider those for whom $\tau = A$ and $o^F = B$. They will optimally choose occupation A if and only if:

$$p(\theta)\beta(1 + a)y \geq p(\theta)(1 + \tilde{n})\beta y, \tag{3}$$

where a higher wage in occupation A, namely $(\beta(1 + a)y)$, more than compensates for the higher job–finding rate in occupation B, namely $(p(\theta)(1 + \tilde{n}))$. It is easy to see that, under Assumption A, Condition (3) holds only for those workers endowed with a small network. All the others (whose mass is μ) optimally decide not to pursue their productive advantage.

We now look at the total probability that type A chooses occupation A:

$$P(o = A|\tau = A) = \Psi(m) + (1 - \mu)(1 - \Psi(m)), \tag{4}$$

where we have rewritten the total probability that type A chooses occupation A as a weighted average of the probability of him choosing occupation A when he has a tradeoff and when he does not. In the first case, he chooses A with probability 1 and in the second case, with

[12]In the analysis, we will frequently use the Law of Total Probability: $P(A|C) = \sum_n P(A|B_n)P(B_n|C)$, where $\{B_n : n = 1, 2, 3, ...\}$ is a partition of the sample space.

probability $(1-\mu)$. The expression in (2) must equal the share of old workers of type A in occupation A (which we previously defined as m) in steady state. Equating the two yields:

$$m^* = \frac{1-\mu\rho}{1+\mu(1-2\rho)}. \tag{5}$$

Lemma 1. *The degree of sorting of workers according to their productive advantage, m^*, is decreasing in μ and increasing in ρ.*

Proof. See Appendix. □

There are two extreme cases: $m^* = 1$ when $\mu = 0$ (no one's choice depends on contacts) and $m^* = \frac{1}{2}$ when $\mu = 1$ (everyone follows in their father's footsteps, regardless of their type). In general, m^* is decreasing in μ and increasing in ρ. The former effect is straightforward to demonstrate: the higher is μ, the more workers will choose purely according to contacts, which availability needs not be aligned with productivity. In contrast, the reason for the latter effect is that, as ρ increases, the pool of workers who choose the same occupation as their father (which is the pool of workers using social contacts) increasingly coincides with the pool of workers who would have chosen that occupation even without social contacts. In this sense, an increase in the persistence of productive types across generations attenuates the negative effects of the use of contacts on worker sorting. As a result, the allocation of the workforce improves. In the limit case when $\rho = 1$, the allocation of workers is no longer affected by μ and there is no mismatch in the economy.

It is useful to derive an expression for the equilibrium value of Ψ, by substituting equation (5) into (2):

$$\Psi^* = \frac{\rho+\mu-2\rho\mu}{1+\mu(1-2\rho)}. \tag{6}$$

Lemma 2. *The share of workers who do not face any tradeoff in their occupational choice, Ψ^*, is decreasing in μ and increasing in ρ.*

Proof. See Appendix. □

An increase in μ, in equilibrium, triggers an increase in the share of workers who choose their occupation according to the availability of contacts. In turn, this lowers Ψ, as more young workers will face a tradeoff. Conversely, an increase in ρ improves the alignment of contacts and comparative advantage, by means of an increase in m^*.

In order to close the model, we now consider the entry of firms. Free entry implies that firms will post vacancies up to the point at which the expected profits from doing so are exactly equal to the fixed cost κ. In order to evaluate profits, firms need to form an expectation of worker productivity in the market. Define $\gamma(m)$ as the probability with which, upon matching, the firm meets with a high–productivity type. This probability will differ from m, since different workers exert different efficiency units of search.

Define the total efficiency units of search as U:

$$U(m) = \Psi(m)[(1-\mu)(1+n) + \mu(1+N)] + (1-\Psi(m))[(1-\mu) + \mu(1+N)] = \\ = 1 + \mu N + \Psi(m)(1-\mu)n. \tag{7}$$

The previous expression is composed of four terms. The former two represent the efficiency units of search exerted by those workers who do not face a tradeoff and can therefore exploit their network, which can be either small or large. The latter two represent the units of search exerted by those with a tradeoff. A proportion $(1-\mu)$ of them decide to stay in the occupation in which they have a comparative advantage, thus giving up their contacts. The remaining share μ decide not to pursue their comparative advantage and to exploit their large network of contacts. This latter group is the only one for which the type and the occupation chosen in equilibrium are not aligned.

Thus, a firm meets with high–productivity type with probability:

$$\gamma(m) = 1 - \frac{(1-\Psi(m))\mu(1+N)}{U(m)}. \tag{8}$$

This expression is in fact the relevant dimension of *mismatch* for aggregate outcomes, that is how poor the allocation of workers is perceived to be by firms. In general, it can be shown that $\gamma(m) < m$ for $\mu > 0$. This is because mismatched workers are overrepresented in the unemployment pool, thus displacing more productive workers. It turns out that an increase in μ has a negative effect on $\gamma(m)$ for two reasons, since not only does it increase the number of mismatched workers, it also increases their probability of being drawn by a firm. On the other hand, ρ has a positive effect on $\gamma(m)$: it reduces the number of those workers with a tradeoff (by raising $\Psi(m)$), and it raises the total efficiency units of search (through the effect on $\Psi(m)$, see Equation 7).

Lemma 3. *The probability of meeting with a productive type in the pool of the unemployed, γ^*, is decreasing in μ and increasing in ρ.*

Proof. See Appendix. □

We now look at the free entry condition. In equilibrium, the expected profits of firms must equate the fixed cost of posting a vacancy:

$$q(\theta)(1-\beta)(1+\gamma(m)a) = \kappa. \tag{9}$$

This shows that equilibrium tightness closely depends on $\gamma(m)$, the likelihood of a firm matching with a high–productivity type. This implies that any factor which affects $\gamma(m)$, will also affect equilibrium tightness in the same direction. This reflects the fact that firms react to changes in the average labor productivity of the economy. For instance, a decrease in $\gamma(m)$ reduces the expected profits from a match and therefore reduces incentives for firm entry. As a result, equilibrium tightness declines. In other words, the fact that some workers decide not to pursue their comparative advantage generates externalities on the demand side, such that aggregate variables are eventually also affected.

Lemma 4. *The equilibrium labor market tightness, θ^*, is decreasing in μ and increasing in ρ.*

Proof. See Appendix. □

2.2 Occupational Followers vs. Occupational Movers

In this section, we derive empirical predictions for the workers who decide to enter the same occupation as their fathers.

We first derive an expression for the degree of occupational persistence across generations in the economy. Such an expression will be useful in the calculations that follow:

$$\begin{aligned} P(o = A|o^F = A) = & P(o = A|o^F = A; \tau = A)P(\tau = A|o^F = A) + \\ & P(o = A|o^F = A; \tau = B)P(\tau = B|o^F = A) = \\ & = \Psi(m) + \mu(1 - \Psi(m)) \equiv \mathcal{P}(m). \end{aligned} \tag{10}$$

We denote as $\mathcal{P}(m)$ the degree of intergenerational occupational persistence, that is, the probability of a young worker choosing the same occupation as his father. We define π as an individual dummy that takes the value 1 if the worker is in the same occupation as his father and 0 otherwise. The model allows us to derive predictions for *occupational followers* (for whom $\pi = 1$) and *occupational movers* (for whom $\pi = 0$). We first consider the probability of finding a job, which is equal to the product of $p(\theta)$ and the efficiency units of search exerted, denoted by s. Occupational followers can exploit their social contacts and therefore will have higher units of search than occupational movers. Thus,

$$E(s|\pi = 1) = m^\pi(1 + N) + (1 - m^\pi)[\mu(1 + N) + (1 - \mu)(1 + n)] > 1 = E(s|\pi = 0). \tag{11}$$

Note that, among occupational followers, some workers will be mismatched (their mass is m^π, an endogenous object that we solve for below). These workers necessarily must have exploited large networks, and thus their efficiency units of search are $(1+N)$ with probability 1; otherwise, they would not have given up their comparative advantage. On the other hand, the non–mismatched workers may have either large or small networks, with probability μ and $(1-\mu)$ respectively. In any case, each subgroup's total efficiency units of search exceed 1. As a result, the job–finding rate of occupational followers will be strictly higher than that of occupational movers.

In order to discuss wages in the model, we first need to derive expressions for the probability of being a productive type, for both followers and movers. In other words, we need to solve for m^π:

$$\begin{aligned} m^\pi \equiv P(\tau = A | o = A; o^F = A) =& \frac{P(\tau = A; o = A; o^F = A)}{P(o = A; o^F = A)} = \\ =& \frac{P(o = A | \tau = A; o^F = A) P(\tau = A; o^F = A)}{P(o = A; o^F = A)} = \\ =& \frac{\Psi(m)}{\mathcal{P}(m)} = \frac{\Psi(m)}{\Psi(m) + \mu(1 - \Psi(m))}, \end{aligned} \tag{12}$$

where we have used $\frac{P(\tau=A;o^F=A)}{P(o=A;o^F=A)} = \frac{P(\tau=A|o^F=A)}{P(o=A|o^F=A)}$. We define $(1-m^\pi)$ as the *mismatch among occupational followers*. If $\mu > 0$, then m^π is strictly smaller than one: the very fact that some workers base their occupational choice on contacts induces some mismatch in the economy. It is easy to verify that repeating the same calculations for the occupational movers shows that there is no mismatch in their case, i.e. $P(\tau = A | o = A; o^F = B) = 1$. This is because, by construction, those workers who do not choose the same occupation as their fathers are not using contacts and therefore they necessarily must have based their decision on productive advantage.

Conditional on matching, the probability of meeting with a high–productivity type among occupational followers is as follows:

$$\gamma^\pi(m^\pi) = 1 - \frac{(1 - m^\pi)(1 + N)\mathcal{P}(m)}{U^\pi(m^\pi)}, \tag{13}$$

where $U^\pi(m^\pi)$ is the total efficiency units exerted by occupational followers:

$$U^\pi(m^\pi) = \mathcal{P}(m)[m^\pi(\mu(1 + N) + (1 - \mu)(1 + n)) + (1 - m^\pi)(1 + N)].$$

This equation consists of three terms: the first two represent the efficiency units of search exerted by the occupational followers who are not mismatched and are using either a large or small network, while the last term, in contrast, represents the mismatched workers, who are necessarily using a large network.

We can now derive the expected wages for the two groups:

$$E(w|\pi = 1) = \beta(1 + \gamma^{\pi}(m^{\pi})a)y < \beta(1 + a)y = E(w|\pi = 0). \tag{14}$$

For $\mu > 0$, we have $\gamma^{\pi}(m^{\pi}) < 1$, and therefore the wages of occupational movers are, on average, strictly higher than the wages of occupational followers. The two predictions derived so far are summarized in the following statement.

Empirical Prediction 1. *Workers who choose the same occupation as their father have, on average, higher job–finding rates and lower wages. Moreover, the former effect is predicted to survive the inclusion of controls for the fixed productive type, while the latter is not:*

1. $E(JF|o = o^F) > E(JF|o \neq o^F)$.
2. $E(w|o = o^F) < E(w|o \neq o^F)$.
3. $E(JF|o = o^F; o = \tilde{o}; \tau = \tilde{\tau}) > E(JF|o \neq o^F; o = \tilde{o}; \tau = \tilde{\tau})$ *with* $\tilde{o}, \tilde{\tau} \in \{A, B\}$.
4. $E(w|o = o^F; o = \tilde{o}; \tau = \tilde{\tau}) = E(w|o \neq o^F; o = \tilde{o}; \tau = \tilde{\tau})$ *with* $\tilde{o}, \tilde{\tau} \in \{A, B\}$.

The intuition behind this result is straightforward. The presence of the father in the same occupation increases the son's available amount of social contacts, thus increasing his job–finding rate. This occurs regardless of the worker's type or occupation. On the other hand, the differences in wages are driven only by selection of workers. In particular, all mismatched workers happen to be occupational followers. However, once we control for the occupational choice and the fixed type (that is, his productivity level), wages are no longer affected by the father's presence.

Another key prediction of the model is that wages are correlated across generations. In this economy, only two wage levels are offered in equilibrium: $\beta(1 + a)y$ to the non–mismatched workers and βy to the mismatched workers. We show that having a mismatched father increases the probability that the son will be mismatched as well. Define $\bar{w}$ as a dummy that takes the value 1 if the worker is non–mismatched (that is, he is a potential high earner), and 0 otherwise. The probability of being a high earner, conditional on having a high–wage father, is as follows (the derivation is reported in the Appendix):

$$P(\bar{w} = 1|\bar{w}^F = 1) = \rho + (1 - \mu)(1 - \rho). \tag{15}$$

Symmetrically, one can also show that:

$$P(\bar{w} = 1|\bar{w}^F = 0) = (1 - \rho) + (1 - \mu)\rho. \tag{16}$$

It is easy to verify that $P(\bar{w} = 1|\bar{w}^F = 1) > P(\bar{w} = 1|\bar{w}^F = 0)$ for $\mu > 0$. In words, having a high–earning father increases the chances of being a high–earning worker. As a result, wages are correlated across generations.

Empirical Prediction 2. *Wages are correlated across generations. Sons of high–earning (low–earning) fathers are more likely to be high earners (low earners) themselves.*

Remarkably, this result does not hinge on any transmission of efficiency level (i.e., unidimensional productivity) across generations, as is usually assumed in the literature. In our case, the transmission of wages across generations is a byproduct of the *transmission of mismatch* across generations.

2.3 Occupational Persistence and Mismatch

In this section, we examine how changes in the structural parameters of the economy (ρ and μ) affect occupational persistence across generations. We also shed light on the relationship between occupational persistence and productive mismatch.

In this economy, occupational persistence is brought about through two different channels: use of contacts and transmission of type (comparative advantage). Strengthening either of these two channels (by increasing μ or ρ, respectively) increases persistence.

Proposition 1. *The probability that a young worker chooses his father's occupation, $\mathcal{P}^*$, is strictly increasing in both μ and ρ.*

Proof. See Appendix. □

In order to understand the comparative statics exercise behind Proposition 1, let us first write an expression for the equilibrium level of persistence, evaluating equation (10) at equilibrium:

$$\mathcal{P}(m^*) = \mu + (1 - \mu)\Psi(m^*). \tag{17}$$

An increase in μ has a twofold effect on persistence: on the one hand, it increases the share of workers who base their occupational choice on contacts; on the other hand, it decreases the share of workers who do not face any tradeoff. These two effects work in opposite directions, but it turns out that the former is always stronger than the latter.

In constrast, an increase in ρ affects persistence only by way of Ψ^*. In particular, a higher probability of transmission of type improves the overall allocation of workers and therefore reduces the probability of facing a tradeoff. Thus, occupational persistence also increases.

We now look more in depth at the relationship between occupational persistence and mismatch. We derive an equation relating the overall level of mismatch in the economy to the mismatch among occupational followers and the degree of occupational persistence:

$$\begin{aligned} m =& P(\tau = A | o = A; o^F = A)\mathcal{P}(m) + P(\tau = A | o = A; o^F = B)(1 - \mathcal{P}(m)) = \\ =& m^{\pi}\mathcal{P}(m) + (1 - \mathcal{P}(m)). \end{aligned} \tag{18}$$

Thus, overall mismatch can be rewritten as a weighted average of the mismatch among occupational followers and that among occupational movers. Plugging equation (10) into (18) and evaluating it in equilibrium yields:

$$m^* = \Psi(m^*) + (1 - \mathcal{P}(m^*)). \tag{19}$$

An implication of (19) is that $(1 - m^*) \leq \mathcal{P}(m^*)$, i.e. the degree of mismatch in the economy is bounded above by the degree of occupational persistence. This is due to the fact that only occupational followers can be mismatched in equilibrium.[13]

2.4 Constrained Efficiency

We now turn to the efficiency properties of our model economy. Inefficiency of the Search Equilibrium (SE hereafter) arises in our setup for two distinct reasons: First, the economy suffers from the usual inefficiency typical of the random search framework. Secondly, the equilibrium level of mismatch does not need to correspond to the efficient level. These two sources of inefficiency are independent of one another, and we describe them below sequentially. We start with the latter, since it is a characteristic feature of our setup, while the former applies to many other search models.

2.4.1 Optimal Level of Mismatch and Occupational Persistence

There are three reasons why the occupational choices of workers (and therefore, the level of productive mismatch) may not be aligned to those of a social planner (SP hereafter): a) workers do not internalize the detrimental effects of the level of mismatch on the equilibrium tightness; b) workers do not internalize that the mismatch has *dynamic* effects via the intergenerational distribution, since the shares of workers with/without a tradeoff (Ψ, in the

[13] This depends on the fact that if a worker is mismatched, it is because he is using social contacts, and this is something that only occupational followers can do. If we were to relax the assumption that social contacts only work through parents, this would no longer be the case. For instance, we could allow workers to have social contacts in the occupation that is not their father's. As long as the probability of having social contacts is higher in their father's occupation, all of our results would still hold.

notation of the SE) depend on the level of mismatch; and c) workers do not internalize the vacancy costs that have to be paid to transform search efficiency units into actual matches.

The externalities described in (a) are not faced by the SP, since he is not constrained by the free entry condition, and can therefore fix the preferred levels of mismatch and tightness independently.[14] Hence, we specify the SP's problem without any reference to this channel, so that the tightness θ is treated as if it were an exogenous parameter, with the understanding that the SP can also operate on this margin.

Since the occupational choice of each subgroup of workers (i.e., combinations of productive types, large/small networks, mismatched/non–mismatched father) affects the equilibrium allocation differently, we allow the SP to choose different degrees of mismatch across groups. Thus, beyond the general level of sorting m, we also define $\{m_1, ...m_8\} \in [0,1]^8$ as the sorting of each subgroup. For instance, m_1 represents the share of workers sorted along their comparative advantage among those with a well–sorted father, the same productive type as their father and a large network (see Appendix for a complete description of the groups). The welfare function to be maximized by the SP is as follows:

$$\begin{aligned} \mathcal{W}^{\mathcal{SP}}(m_1, ..., m_8) = 2\Big\{ & m\rho\mu[p(\theta)y(1+m_1 a) - \kappa\theta](1+m_1 N) \\ + \quad & m\rho(1-\mu)[p(\theta)y(1+m_2 a) - \kappa\theta](1+m_2 n) \\ + \quad & m(1-\rho)\mu[p(\theta)y(1+m_3 a) - \kappa\theta](1+(1-m_3)N) \\ + \quad & m(1-\rho)(1-\mu)[p(\theta)y(1+m_4 a) - \kappa\theta](1+(1-m_4)n) \\ + \quad & (1-m)\rho\mu[p(\theta)y(1+m_5 a) - \kappa\theta](1+(1-m_5)N) \\ + \quad & (1-m)\rho(1-\mu)[p(\theta)y(1+m_6 a) - \kappa\theta](1+(1-m_6)n) \\ + \quad & (1-m)(1-\rho)\mu[p(\theta)y(1+m_7 a) - \kappa\theta](1+m_7 N) \\ + \quad & (1-m)(1-\rho)(1-\mu)[p(\theta)y(1+m_8 a) - \kappa\theta](1+m_8 n)\Big\}, \end{aligned} \tag{20}$$

where m, which is the total share of workers who realize their comparative advantage, is a strictly increasing function of $\{m_1, ...m_8\}$ (see Appendix). It is straightforward to show that the SP indeed cares about the consequences of occupational allocation on future generations, via intergenerational transmission. In other words, he internalizes the fact that the share of each subgroup is a function of m (channel (b) described above).[15]

[14] It is true that, if workers were aware of the fact that equilibrium tightness depends on the level of mismatch, there would be less mismatch. However, at the same time, this margin is simply ignored by the SP. As a consequence, (a) does not represent a reason for the SP to produce less mismatch than in the SE.

[15] For instance, the share of subgroup 1 (whose allocation is described by the m_1 variable) is a function of the mismatch of the parent generation, as well as of the transmission parameters ($m\rho\mu$). In Equation 20, we are considering a steady–state allocation (in which the mismatch of both generations is the same).

The total derivative of $\mathcal{W}^{\mathcal{SP}}$ w.r.t. the sorting of any of these groups is composed of two different terms:

$$\frac{d\mathcal{W}^{\mathcal{SP}}}{dm_j} = \underbrace{\frac{\partial\mathcal{W}^{\mathcal{SP}}}{\partial m}\frac{\partial m}{\partial m_j}}_{reallocation\ effect} + \underbrace{\frac{\partial\mathcal{W}^{\mathcal{SP}}}{\partial m_j}}_{inner\ effect} \qquad \forall\, j \in \{1, ..., 8\}. \tag{21}$$

The *reallocation effect* has to do with the intergenerational distribution. In other words, the SP understands that the equilibrium shares depend on the allocation. The *inner effect*, though it involves a similar tradeoff as the one faced by workers, includes the additional consideration of the vacancy costs. In general, both of these effects depend on all the elements of the vector $\vec{m} = \{m_1, ...m_8\}$ and it is not easy to fully characterize the SP's equilibrium. Instead, we derive conditions under which the SE allocation is not efficient (see the Appendix for details).

Define $N^* = \frac{N}{1+N} - \frac{N}{1+N}\frac{\kappa\theta}{p(\theta)y}$ and $n^* = \frac{n}{1-n} - \frac{n}{1-n}\frac{\kappa\theta}{p(\theta)y}$.

Proposition 2. *Depending on the parameter values, the SE allocation $\vec{m}^{SE}$ does not necessarily coincide with the efficient one.*

- *Under $a > N^*$ and $a > n^*$:*
  The SP wants to achieve less mismatch than in the SE.
- *Under $a < N^*$ and $a < n^*$, $\exists\bar{\rho} > \frac{1}{2}$ such that $\forall\rho < \bar{\rho}$:*
  The SP wants to achieve more mismatch than in the SE.

Proof. See the Appendix. □

Intuitively, high values of N and n, along with low values of a and κ, and ρ arbitrarily close to $\frac{1}{2}$, can drive the SP to choose even more mismatch than in the SE. In this case, the benefits from using better search technology more than offset the costs (i.e. a drop in labor productivity) for the economy as a whole.

Another implication of Proposition 2 is that, in general, the level of occupational persistence in the SE is not socially optimal.[16]

2.4.2 Optimal Level of Tightness

We now turn to the optimal level of labor market tightness in the economy. We define $\mathcal{W}^{\mathcal{SE}}$ as the welfare achieved by the SE, which corresponds to aggregate net income, that is,

[16] Occupational persistence is a function of the allocation of workers $\{m_1, ...m_8\}$ (see the Appendix for details).

the difference between expected match output and the vacancy cost multiplied by the total number of search efficiency units in the economy:

$$\mathcal{W}^{\mathcal{SE}} = 2[p(\theta)(1+\gamma a)y - \kappa\theta]U. \tag{22}$$

We take the level of γ and U, which are endogenous variables that depend on the level of m, as given. It is easy to show that the equilibrium level of θ is generally inefficient, unless we are in a knife–edge case. In particular, the derivative of welfare with respect to market tightness is as follows:

$$\frac{\partial \mathcal{W}^{\mathcal{SE}}}{\partial \theta} = 2\kappa \left[\frac{\eta_V(\theta) - (1-\beta)}{1-\beta}\right] U. \tag{23}$$

where we have used the fact that $q(\theta)(1+\gamma a)y = \frac{\kappa}{1-\beta}$ and the definition $p(\theta) = q(\theta)\theta$, and where $\eta_V(\theta) = \frac{q'(\theta)\theta + q(\theta)}{q(\theta)}$ is the elasticity of the matching function with respect to V. Generally, $\eta_V(\theta)$ will differ from $(1-\beta)$, thus making the equilibrium inefficient.[17]

2.4.3 Occupational Persistence and Welfare

In this section, we establish the relationship between occupational persistence and welfare. The former is an endogenous object, and therefore we first need to understand how welfare is affected by changes in the determinants of persistence (that is, the parameters μ and ρ). Changes in μ or ρ have, in principle, an ambiguous effect on welfare. More specifically, each of them affects three different variables simultaneously: the total amount of efficiency units of search U, the sorting of workers across occupations γ and the level of tightness θ. The total derivative of welfare with respect to μ is as follows:

$$\frac{d\mathcal{W}}{d\mu} = \frac{\partial \mathcal{W}}{\partial \mu} + \frac{\partial \mathcal{W}}{\partial \theta}\frac{d\theta}{d\mu}. \tag{24}$$

The second term reflects the externalities mentioned in the previous sections, whereby variation in the average level of labor productivity transmits to the equilibrium level of tightness ($\frac{\partial \theta}{\partial \mu} < 0$). Whether this increases (decreases) welfare depends on whether the equilibrium level of tightness is inefficiently low (high).

With respect to the first term (the direct effect), it can itself be decomposed into two different margins:

$$\frac{\partial \mathcal{W}}{\partial \mu} = 2\left[p(\theta)(1+\gamma a)y - \kappa\theta\right]\frac{\partial U}{\partial \mu} + 2p(\theta)ayU\frac{\partial \gamma}{\partial \mu}. \tag{25}$$

[17] This is the same inefficiency studied by Hosios (1990), which arises from a combination of search externalities and Nash bargaining with pre–fixed shares of surplus division.

The first term is always positive, reflecting the improvement in the search technology of the economy ($\frac{\partial U}{\partial \mu} > 0$). The second term, on the other hand, is always negative, reflecting that it is now harder for firms to match with productive types ($\frac{\partial \gamma}{\partial \mu} < 0$). The overall effect is ambiguous, depending on the specific parametrization.

Turning to the effect of ρ, the same decomposition can be performed:

$$\frac{d\mathcal{W}}{d\rho} = \frac{\partial \mathcal{W}}{\partial \mu} + \frac{\partial \mathcal{W}}{\partial \theta}\frac{d\theta}{d\rho}. \tag{26}$$

The aforementied considerations apply here as well. One important difference is that here the effect of ρ on market tightness is positive ($\frac{\partial \theta}{\partial \rho} > 0$). Moreover, the direct effect (i.e. the first term) is now unambiguously positive:

$$\frac{\partial \mathcal{W}}{\partial \rho} = 2\left[p(\theta)(1+\gamma a)y - \kappa\theta\right]\frac{\partial U}{\partial \rho} + 2p(\theta)ayU\frac{\partial \gamma}{\partial \rho}. \tag{27}$$

Not only does an increase in ρ inflate the efficiency units of search (since $\frac{\partial U}{\partial \rho} > 0$) through an increase in occupational persistence, but it also enhances the sorting of workers ($\frac{\partial \gamma}{\partial \rho} > 0$).

An implication of the discussion so far is that, in general, there is no one–to–one relationship between occupational persistence and welfare.

Proposition 3. *Changes in the degree of occupational persistence, $\mathcal{P}$, can be associated with either an increase or a decrease in the level of welfare, $\mathcal{W}$. If $\eta_V(\theta) \geq (1-\beta)$ and $\frac{N}{n} \leq \hat{N}$, (where $\hat{N} = f(\mu, \rho, n)$, see Appendix), then:*

- *An increase (decrease) in persistence generated by an increase (decrease) in μ has negative (positive) effects on welfare: $\frac{\partial \mathcal{P}}{\partial \mu} > 0$ and $\frac{\partial \mathcal{W}}{\partial \mu} \leq 0$.*
- *An increase (decrease) in persistence generated by an increase (decrease) in ρ has positive (negative) effects on welfare: $\frac{\partial \mathcal{P}}{\partial \rho} > 0$ and $\frac{\partial \mathcal{W}}{\partial \rho} \geq 0$.*

Proof. See the Appendix. □

In other words, it is crucial to understand the extent to which the different channels are generating persistence. For instance, if the latter is entirely due to the transmission of the productive type, then it is not a sign of a suboptimal allocation of workers into occupations. If, instead, persistence is brought about by workers choosing according to the avilability of their father's contacts, then it may be a signal of underlying mismatch.

In this sense, it is unclear whether occupational persistence across generations is socially desirable, unless we are able to decompose it into its sources. In order to answer this question, in Section 4 we will construct a structural model of intergenerational transmission and occupational choice, so that in Section 5 we will be able to perform a decomposition exercise.

3 Empirical Evidence

In this section, we document the degree of occupational persistence across generations in the UK and test the key predictions (Empirical Prediction 1) of the model developed in the previous section. To this end, we use the British Household Panel Survey (BHPS) and in particular the dataset constructed by Lo Bello & Morchio (2017).

3.1 The Data

The BHPS is a yearly survey taken by about 10,000 individuals per year in the UK. It was first carried out in 1991, and the last available wave for this study is 2008. The survey is characterized by a fairly high follow–up rate, with more than 90% of the individuals being interviewed also in the subsequent year, and a number of new households entering the sample each year. In total, 32,377 individuals were interviewed in the BHPS during the period 1991-2008. We restrict our sample to males[18] aged 16–65, and are left with 12,982 individuals, for a total of 1,023,888 monthly observations. Individuals report a detailed job history of the previous year, including all the employment/unemployment spells, along with several job characteristics of each job (among them, the occupational group). In this way, we are able to construct long labor market histories for each individual (potentially up to 216 months) and, more importantly, we are able to observe transitions at the monthly frequency. Apart from a detailed job history, each individual provides demographic information, including gender, age, education, occupation, race, marital status, region of residence, etc. One key feature of the dataset is that it allows us to connect individuals to their fathers and to track them both over time.

The job–finding rate is defined as the monthly probability of transiting from unemployment to employment. Wages are calculated by dividing the total monthly labor income by the number of hours normally worked per week multiplied by four (the information on hours worked is only available for the current job at the moment of the interview, that is, it is recorded annually).

[18] We exclude women from the sample for several reasons: i) employment rates of men and women are substantially different (especially for the parent generations); ii) in order to maintain comparability to the rest of the literature, which also excludes women; iii) in previous work we found that, although occupational following is also prevalent among women, there is no evidence that mothers serve as network providers (see Lo Bello & Morchio 2017).

3.2 Intergenerational Occupational Persistence

The data allows us to study the extent of occupational persistence across generations. We compute the distribution of workers across occupations and study the probability of a father and his son belonging to the same occupational group. In order to account for the unequal distribution of workers across occupations, we construct likelihood ratios. For the sake of concreteness, we define the *persistence index* $\mathcal{P}_j$ as the ratio between the probability of belonging to a given occupation j conditional on the father also belonging to it and the unconditional probability of belonging to occupation j:

$$\mathcal{P}_j = \frac{P(o = j | o^F = j)}{P(o = j)} .$$

We define nine occupational groups, following the SOC aggregation by major group (the 1–digit level), as established by the Employment Department Group and the Office of Population Censuses and Surveys. Results are shown in Table I:

Occupational group	**Likelihood Ratio**
Managers & Administrators	1.29
Professional	2.60
Associate Professional & Technical	1.62
Clerical & Secretarial	1.26
Craft & Related	1.55
Personal & Protective Service	1.58
Sales	1.34
Plant & Machine	1.94
Agriculture & Elementary	2.67
Average (unweighted)	1.76
Average (weighted)	1.72

Table I. Occupational Persistence Indexes (Likelihood Ratios). Source: BHPS (1991–2008).

We find a large degree of occupational persistence. The estimated likelihood ratios of occupational persistence are greater than 1, indicating that a worker is more likely to belong to a given occupation if his father also belongs to it. The average likelihood ratio is estimated to be 1.76, implying that an individual is 76% more likely to be in a given occupation if his father is as well (this excess probability ranges from 29% to 167%, depending on occupation). Interestingly, persistence does not appear to vary systematically with the occupation's skill level or wage.[19] Repeating the same exercise at the 2–digit level, we find that the average

[19]This seems to suggest that borrowing constraints are not playing a major role in occupational choice. We also checked whether likelihood ratios vary by father's income within an occupational group. Our results

unweighted and weighted likelihood ratios are 5.69 and 4.71, respectively (see Table XIII in the Appendix).

Part of the persistence might be explained by the usual socioeconomic variables, such as age, education or the region of residence. In order to account for this, we estimate linear probability models which regress the probability of belonging to a given occupation (as opposed to any other) on a number of covariates. We estimate the model for each occupation, and present the results in Table XII (in the Appendix). The estimates reveal that a worker is *ceteris paribus* between 1.59 and 15.1 p.p. more likely to belong to the same occupation as his father. These are large probability differences and are highly statistically significant in all of the 9 occupations. Interestingly, the average conditional likelihood ratio implied by these estimates is actually 10% higher than the unconditional one. This suggests that covariates do not play much of a role in explaining occupational persistence.

We also perform some addditional checks in order to investigate whether occupational persistence is primarily related to regional factors. For instance, living in a poor region (or a region with a limited variety of job opportunities) might mechanically increase the likelihood ratio. In that case, we would not be comparing the conditional probability to the *correct* unconditional probability. We plot the region–specific weighted average likelihood ratio against the average regional wage and a measure of occupational concentration (the Herfindahl index), and find that neither of these two dimensions can predict persistence (results are shown in Figures 5 and 6 in the Appendix,). This provides reassurance that regional factors are not playing a major role in determining the results.[20]

3.3 Intergenerational Occupational Persistence and Occupational Attachment

Thus far, we have shown that a worker's occupational affiliation tends to be highly correlated with that of his father. In this section, we investigate whether this phenomenon is persistent over the life–cycle. This is important because young workers, who are potentially sampling different occupations, may be those who are driving the likelihood ratios estimated in the previous section. More importantly, these young workers might be using their father's occupation as a stepping stone to their eventual occupation (possibly to avoid unemployment). If this is the case, then occupational persistence would be a short–run phenomenon, with

show that occupational persistence decreases only slightly. For instance, if we only consider the top 1/3 of earners in each of the occupations, the average likelihood ratio drops to 1.61 (Table XIV in the Appendix).

[20]An important caveat is that we only have 19 regions. It is plausible that the relevant level for the father–son transmission is finer than that. However, the sample size does not allow us to estimate occupation–specific indexes of persistence at the regional level.

limited consequences for the allocation of workers to occupations.

First, we document that likelihood ratios are not decreasing over the life–cycle. For instance, the average (unweighted) likelihood ratio is 1.88 for workers younger than 20, as opposed to 1.91 for workers aged 25–30 and 2.03 for workers older than 30 (see Table XV in the Appendix).

Second, we look at the length of the occupational spells of followers, as opposed to those of movers. The average occupational tenure is 2.16 years for followers, and 1.73 years for movers.[21] This is true also for occupations chosen very early in an individual's career. Figure 1 plots the share of workers still in their first occupation against the number of years of labor market experience, for followers and movers separately.

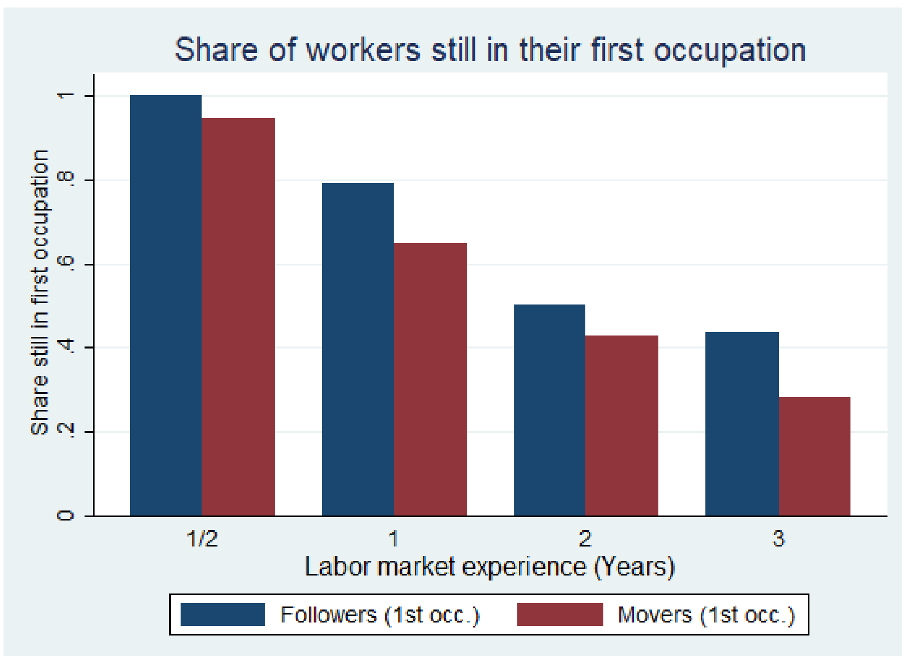

Figure 1. Share of workers still in their first occupation, by years of labor market experience. Source: BHPS (1991–2008).

We can see that a worker who starts his career in the same occupation as his father's is substantially less likely to exhibit occupational mobility. For instance, after two years from the start of their first employment spell, 60% of occupational followers will not have changed occupation[22] as compared to 49% of occupational movers. At the same time, these statistics reveal a large degree of *hysteresis* in occupational transitions. In other words, the initial

[21]When we restrict our attention only to the spells that we observe from the start, we find again that followers tend to be more attached to their occupation (average tenure of 1.84 years versus 1.69 years).

[22]In Figure 1, we do not count flows back into the original occupation as *still in the same occupation*. If we were to do that, we would find a slightly larger difference between followers and movers.

occupation is a good predictor of occupational affiliation even several years after the start of the employment spell. In this sense, the father's influence on the initial occupational choice may have long–lasting consequences for his son's outcomes and the aggregate allocation.

As an additional piece of evidence, we look at whether the contemporaneous presence of the father in the same occupation has an impact on the probability of changing occupation. To this end, we run the following regression:

$$OC_{i,t} = \alpha + \beta\pi_{i,t-1} + \boldsymbol{\gamma}\boldsymbol{X}_{i,t} + \epsilon_{i,t}, \tag{28}$$

where $OC_{i,t}$ is a dummy taking the value 1 if the occupation at time t is different from the one at $t-1$ (i.e. there has been an occupational switch[23]) and 0 otherwise; $\pi_{i,t}$ is a dummy variable that takes the value 1 if the occupation of son i and his father coincide at time t, and 0 otherwise; $\boldsymbol{X}_{i,t}$ is a vector of control variables that include a third degree age polynomial, dummies for educational categories and occupational groups (observed for the employed, imputed for the unemployed), marital status, ethnic group, smoking behavior (to capture health level), region of residence and quarter dummies; $\epsilon_{i,t}$ is the idiosyncratic error term.

We estimate Equation 29 with pooled OLS, random effects and fixed effects, with the estimates of β shown in Table II.

Dependent Variable: Occupational Change

	(1) POLS	(2) RE	(3) FE
Father in same occupation ($\pi_{i,t-1}$)	-0.00811***	-0.00794***	-0.00872***
	(0.002)	(0.002)	(0.002)
Average in–sample OC rate	0.0265	0.0265	0.0265
N	53208	53208	53208
R^2	0.015	-	0.014
Number of pairs	-	938	938

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

Table II. Regressions of Occupational Change (transition from one occupation to another); coefficient for **father in same occupation last month** (dummy variable), standard errors and average occupational change rate in the regression sample. Model 1 is a pooled OLS regression, model 2 is a random effects GLS regression, and model 3 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behavior, marital status, ethnicity, quarter and occupation of employment. Source: BHPS (1991–2008).

[23]The occupational switch can take place either through unemployment (where we compare the previous and subsequent occupations) or not (direct employment–to–employment switch).

We find that if the father is employed in the same occupation, there is a substantial reduction in the likelihood of changing occupation. The estimated impact is in the region of -0.8/0.9 p.p., which represents about one–third of the average in–sample monthly occupational change rate (2.65 p.p.). One possible interpretation is that some workers are more mobile than others in general, and therefore they will happen to be less often in the same occupation as their father, thus mechanically generating a correlation between the two variables. However, notice that: i) we are exploiting the exact timing of the transitions (using the lagged persistence variable), thus making this interpretation less likely; ii) in Column 3, we are controlling for individual fixed effects, ruling out this type of explanation. The estimated coefficient, which is quite stable across specifications, suggests that a worker is more reluctant to leave his father's occupation, even on top of any unobserved fixed heterogeneity.

In the following subsections, we turn to testing the two main predictions of the model, namely that occupational followers[24] have, on average, higher job–finding rates but lower wages.

3.4 Occupational Persistence and Job–Finding Rates

The first prediction of the model is that *occupational followers*, i.e. sons who are in the same occupation as their father, will on average have a higher job–finding rate. Given that we observe employment status and occupational affiliation, we can directly test this prediction. Therefore, we run the following regression:

$$JF_{i,t} = \alpha + \beta\pi_{i,t} + \boldsymbol{\gamma}\boldsymbol{X}_{i,t} + \epsilon_{i,t}, \tag{29}$$

where $JF_{i,t}$ is defined only for the unemployed and takes the value 1 if a job is found at time t and 0 otherwise; $\pi_{i,t}$ is a dummy variable that takes the value 1 if the occupation of son i and his father coincide at time t and 0 otherwise;[25] $\boldsymbol{X}_{i,t}$ is a vector of control variables that include a third–degree age polynomial, dummies for educational categories and occupational groups (observed for the employed, imputed for the unemployed), marital status, ethnic group, smoking behavior (to capture health level), region of residence and quarter dummies; $\epsilon_{i,t}$ is the idiosyncratic error term.

We estimate Equation 29 with pooled OLS, random effects and fixed effects, and present the estimates of β shown in Table III:

[24]Our working definition of *occupational follower* is based on $\pi_{i,t}$. However, the following results for JF rates and wages hold for more stringent definitions of *occupational follower* (for instance, starting the spell as a follower and then staying there for at least 6 or 12 months).

[25]The occupation of an unemployed individual is assumed to be the one in which a job will be found at the end of the unemployment spell. Moreover, this variable is defined only for those with an employed father.

Dependent Variable: Job–Finding Rate

	(1) POLS	(2) RE	(3) FE
Father in same occupation ($\pi_{i,t}$)	0.0546***	0.0531**	0.0546**
	(0.016)	(0.021)	(0.026)
Average in–sample JF	0.125	0.125	0.125
N	4142	4142	4142
R^2	0.057	-	0.046
Number of pairs	-	401	401

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table III. Regressions of Job–Finding Rate (transition from Unemployed to Employed); coefficient for **father in same occupation** (dummy variable), standard errors and average job–finding rate in the regression sample. Model 1 is a pooled OLS regression, model 2 is a random effects GLS regression, and model 3 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behavior, marital status, ethnicity, quarter, and occupation of search/employment. Source: BHPS (1991–2008).

We find that *occupational followers* have, on average, a substantially higher monthly job–finding rate (+5.4 p.p.) relative to *occupational movers*. Given that the unconditional probability of finding a job is estimated to be 12.5%, an individual whose father is in the same occupation increases his monthly probability of finding employment by about 42%. Importantly, the effect is robust to the inclusion of individual fixed effects (column 3), which control for unobserved heterogeneity across individuals. The identification of individual fixed effects is made possible by the panel structure of the data. The coefficient presented in column (3) of Table III is estimated by exploiting the variation in π (i.e. whether or not the father is in the same occupation), *within* the son's working life. This is in line with our model, which predicts that even after controlling for occupation and fixed type (τ), the father's occupation is still an important determinant of the individual's job–finding rate. We also find that these effects are robust to the exclusion of the self–employed from the sample (see Table XVI in the Appendix).

To the extent that social networks are slowly accumulated over time (as will be the case in our quantitative model), we also look at whether the impact of the father's occupation changes with his occupational tenure. Consistent with the prediction of the theory, we obtain a positive (though not statistically significant) coefficient for the interaction between $\pi_{i,t}$ and father's tenure, as shown in Table XVII in the Appendix.

Finally, we examine whether the impact of $\pi_{i,t}$ changes with age. We find that the effect is particularly high (up to +12 p.p.) among the youngest workers and then monotonically declines thereafter (Figure 7 in the Appendix). This piece of evidence lends support to our interpretation: young workers, who lack experience in the market, are expected to depend more heavily on their father's contacts. In a dynamic setting, like that to be developed in Section 4, workers accumulate contacts themselves, and therefore the influence of their father will fade over time.

3.5 Occupational Persistence and Wages

The second key prediction of the model is that *occupational followers* have lower wages on average, which according to the model is entirely due to selection. In fact, upon choosing an occupation, the wage only depends on the individual's productivity level in that occupation. That is, the father's occupational affiliation no longer matters. Thus, the difference between *occupational followers* and *occupational movers* is predicted to disappear after controlling for individual fixed effects (which capture the productivity level).

In order to test this hypothesis, we construct for each individual an index of how long he spent in the same occupation as his father during his working life. We do this for two reasons: first, we want to construct groups that do not change over the life–cycle of individuals; and second, it is generally unclear whether we should look at the father's occupation at the moment of the wage observation or at the start of the job spell.[26] Thus, we define q_i as the fraction of his months in employment that individual i spent in the same occupation as his father:

$$q_i = \frac{\sum_t \pi_{i,t}}{\sum_t E_{i,t}}, \tag{30}$$

where $E_{i,t}$ is a dummy taking the value 1 if the individual is employed in period t and 0 otherwise. The index q_i ranges from 0 to 1, and is a measure of the number of months (out of those in which he was employed) during which his occupation coincided with his father's. In Figure 2, we plot the wage profiles of three groups: those for which $q_i = 0$, $q_i \geq 0.5$ and $q_i = 1$.

We can see that the wages of those who spent more time in the same occupation as their father are lower by up to 20% on average throughout their entire working life. Remarkably, this difference appears to be constant over the lifecycle.

We investigate whether these differences depend on observable heterogeneity across workers. In particular, we estimate the following regressions:

[26] In the Appendix (Figure 8), we show that differences in wages are robust to alternative definitions of *occupational followers*.

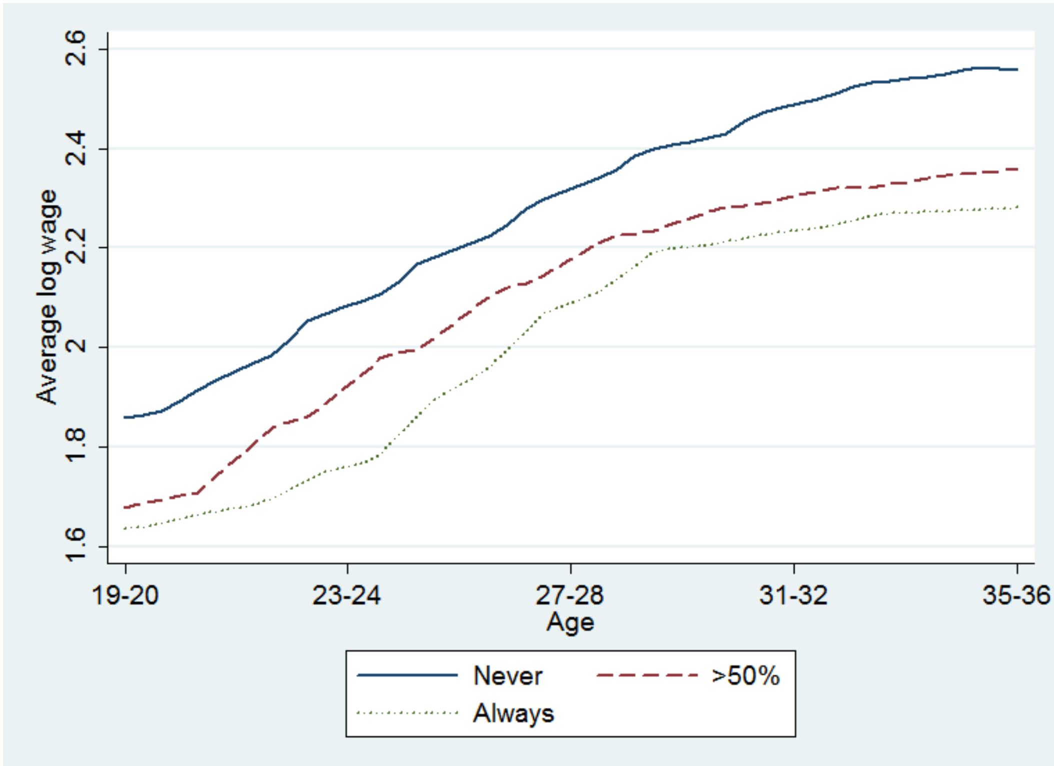

Figure 2. Wage profiles by proportion of employed worklife spent in the same occupation as the father. Source: BHPS (1991–2008).

$$\log(w_{i,t}) = \alpha + \beta q_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \epsilon_{i,t}, \tag{31}$$

where $\log(w_{i,t})$ is the natural logarithm of the wage (observed at the annual frequency); q_i is defined in (30); $\boldsymbol{X}_{i,t}$ is a vector of control variables; and $\epsilon_{i,t}$ is the idiosyncratic error term.

We estimate Equation 31 by POLS and RE. Given that controlling for unobserved heterogeneity is not possible in this regression model, we estimate the same regression again except that we replace the independent variable q_i with $\pi_{i,t}$, the time–specific persistency dummy variable:

$$\log(w_{i,t}) = \alpha + \beta \pi_{i,t} + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \epsilon_{i,t}. \tag{32}$$

The first two columns of Table IV indicate that individuals who have spent more time in the same occupation as their father tend to earn lower wages, even after adding all the controls. We find that those who spend 10% more of their employed working life in the same occupation as their father earn 1.45%–1.52% lower wages, on average. Columns 3 to 5 present the estimates of β in Equation (32). We find an average discount of 7.6% associated with the presence of the father in the same occupation. However, this effect declines to 2.4% (which is barely statistically significant at the 90% confidence level) when we allow for RE in the regression model, and is reduced even further (to the point that it is neither statistically nor economically significant) when individual fixed effects are included in the regression. Overall, these results lend strong support to our theory.[27]

In the Appendix (Table XVIII), we show that these results are robust to trimming the sample (i.e. removing the top and bottom 1% or 5% of the wage observations).

Dependent Variable: Log Hourly Wage

	(1) POLS	(2) RE	(3) POLS	(4) RE	(5) FE
Share of time in same occ. as father (q_i)	-0.145***	-0.152***			
	(0.017)	(0.035)			
Father in same occ. ($\pi_{i,t}$)			-0.076***	-0.024*	-0.0003
			(0.014)	(0.013)	(0.014)
N	6485	6485	4776	4776	4776
R^2	0.623	-	0.604	-	0.624
Number of pairs	-	922	-	850	850

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

Table IV. Regressions of Log Hourly Wage; coefficient for **share of time spent in same occupation as father** (from 0 to 1), standard errors and **father in same occupation** (dummy variable). Models 1 and 3 are pooled OLS regressions, models 2 and 4 are random effects GLS regressions, and model 5 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education and occupation, second–order polynomials in occupational tenure and potential labor market experience, firm size, region of residence, smoking behavior, marital status, ethnicity, and year. Source: BHPS (1991–2008).

3.6 Unemployment Risk and Wages

We have so far established that occupational followers: i) tend to spend less time in unemployment; and ii) tend to earn lower wages. However suggestive, these two pieces of evidence *per se* do not imply that individuals actually face the tradeoff (between employment prospects and wages) described in our model. For instance, it could be that these two observations are the results of looking at two different subsamples (the unemployed and the employed), which may differ in other characteristics as well.

In order to overcome this issue, we exploit the entire working life of the workers in the

[27]Another natural test of our theory would be to look at the wages of occupational switchers. Inutitively, our model predicts that wages should decline (increase) when individuals move to (from) their father's occupation. Unfortunately, due to the limited sample size, we cannot test this implication.

sample. For each worker i, we compute the share of time spent employed $\bar{E}_i = \frac{\sum_t E_{i,t}}{\sum_t E_{i,t}+U_{i,t}}$ (a measure of his employment prospects) and the average monthly wage[28] earned throughout his working life $\bar{W}_i$ (a measure of lifetime labor earnings). In order to compute these lifetime statistics, we include observations from age 25 onwards.[29]

We find that $\bar{E}_i$ is positively related to q_i, while the opposite is true for $\bar{W}_i$ (Figure 3). Occupational followers appear to be characterized by better employment prospects and lower wages.

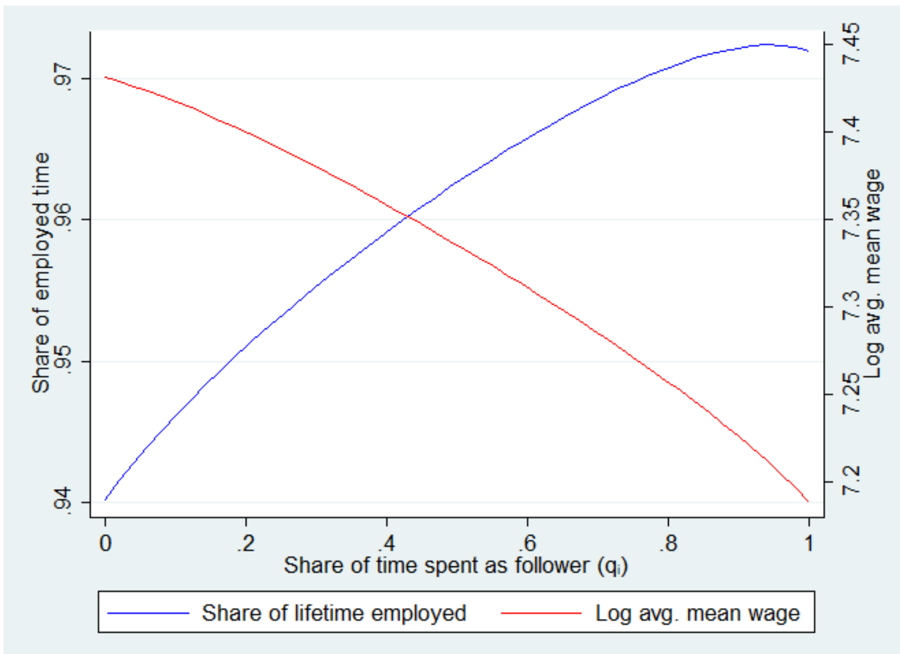

Figure 3. Locally weighted linear polynomial regression (degree 1, bandwidth 0.5) of share of lifetime employed and log average mean wage against the share of time spent as a follower. Source: BHPS (1991–2008).

In Table V, we show the residual partial correlation between the aforementioned variables, controlling for fixed characteristics of individuals (i.e. education and race). Column 1 and 2 deliver the same message as Figure 3: occupational followers tend to have lower wages (by up to -24%) but better employment prospects (on average, they are employed for 5.2 p.p. more of their total time spent in the labor force). Interestingly, employment prospects and wages are generally positively correlated (Column 3), but their respective correlations with

[28]This measure incorporates the unemployment risk margin as well. We construct it in the following manner: first, for each year, we multiply the monthly wage by the number of months that the individual is employed; we then sum them over the years; and finally, we divide the total by the number of wage observations (to correct for the unbalanced nature of the panel).

[29]The rationale behind this is to ensure that we are not capturing effects related to variation in the age of entry into the labor market.

q_i have opposite signs. The sign of both of these correlations is robust to the introduction of the other control variables. In other words, conditional on lifetime employment prospects, followers tend to have lower wages (Column 4); and conditional on the average lifetime wage, tend to spend more time employed (Column 5).

	Dependent Variable				
	(1)	(2)	(3)	(4)	(5)
	$\bar{W}_i$	$\bar{E}_i$	$\bar{E}_i$	$\bar{W}_i$	$\bar{E}_i$
Share of time in same occ. as father (q_i)	-0.239***	0.0526**		-0.290***	0.0514***
	(0.069)	(0.025)		(0.061)	(0.014)
Log avg. mean wage ($\bar{W}_i$)			0.103***		0.108***
			(0.009)		(0.009)
Share of litefime employed ($\bar{E}_i$)				2.006***	
				(0.169)	
Controls (educ, race)	X	X	X	X	X
N	524	601	524	524	524
R^2	0.064	0.014	0.209	0.267	0.228

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

Table V. Regressions of log average mean wage ($\bar{W}_i$) and share of lifetime spent employed ($\bar{E}_i$); coefficient for **share of time spent in same occupation as father** (from 0 to 1), log average mean wage and share of lifetime spent employed; standard errors in parentheses. All models include dummies for education and ethnicity. Source: BHPS (1991–2008).

Overall, the empirical evidence presented here is consistent with the tradeoff featured in our simple model, strongly suggesting that workers indeed face such a tradeoff.

3.7 Absolute vs. Comparative Advantage

One of the key findings presented above, namely the wage discount of *occupational followers*, may also be consistent with theories of occupational sorting based on absolute advantage, such as Groes *et al.* (2014). According to this alternative view, sons of high–wage fathers tend to be high–ability workers themselves and therefore they may be more prone to change occupation (perhaps because they face lower switching costs or because they have a higher level of talent to realize). If this were indeed the case, then the wage discount of occupational followers would be delivered by a mechanism that does not imply any occupational misallocation.

We argue that such an alternative view implies a negative relationship between the father's wage and the son's likelihood of being a follower. Our theory, which is based on selection along the comparative advantage margin, implies exactly the opposite relationship: there is a higher chance that sons of high-wage (and therefore well-matched according to our theory) fathers are more likely to be followers. This is due to the fact that the sons of high–wage fathers face less of a tradeoff in their occupational choice than those of low–wage fathers.

Furthermore, the selection mechanism based on absolute advantage implies that, once we control for the individual's wage, the father's wage does not have any residual predictive power for persistence (to the extent that the individual's wage accurately captures the individual's ability level).[30] In contrast, according to our theory, the father's wage maintains its positive predictive power. To see this, notice that the probability of being a follower conditional on the individual's wage is as follows:

$$\begin{aligned}
P[\pi = 1|\bar{w} = 0] =& 1, \\
P[\pi = 1|\bar{w} = 1] =& P[\bar{w}^F = 1|\bar{w} = 1]\,\rho \\
&+ P[\bar{w}^F = 0|\bar{w} = 1]\,(1-\rho).
\end{aligned}$$

Given that $\rho > 1/2$, this conditional probability is positively correlated with the level of the father's wage. The intuition behind this is straightforward: high–wage (i.e. well–sorted) sons are occupational followers to a larger extent if their father is also high–wage (i.e. well–sorted).

We test the opposite predictions of the two theories by regressing the likelihood of being a follower on the father's wage, both unconditional and conditional on the individual's wage (Columns 1 and 2 of Table XIX in the Appendix). In both cases, the strongly positive correlation between the two variables is supportive of our theory, which is based on comparative advantage. To the extent that the level of productivity is better captured by the average (lifetime) log wage, we repeat the regression using these average measures for both the father's and the son's wage. As shown in columns 3 and 4 of Table XIX in the Appendix, results are nearly unchanged.

4 The Quantitative Model

In this section, we develop a quantitative model that extends the model in Section 2 in a number of dimensions. First, the model presented here is dynamic, such that individuals face a stochastic ageing process. Second, we have O different occupations. Third, we in-

[30] If we were to consider measurement error in wages, the father's wage would still retain negative predictive power.

troduce non–pecuniary benefits (preferences) for occupations,[31] which are composed of a permanent as well as a temporary component. The permanent component is allowed to be correlated across generations, providing an additional source of occupational persistence. In constrast, the introduction of shocks to non–pecuniary benefits allows the model to generate occupational mobility over the life–cycle.[32] Fourth, we introduce occupation–specific human capital. Fifth, we add endogenous job separations. We retain the assumption that social contacts are occupation–specific, and that fathers help their sons find a job in their occupation, without internalizing that in their own occupational choice. Both social contacts and occupation–specific human capital evolve over time.

4.1 The Model Environment

Time is discrete ($t = 0, 1, 2, ...$) and goes on forever. The economy is divided into a discrete number of submarkets O, which represent the different occupations. A measure 2 of workers populate the economy. All agents (both workers and firms) are risk–neutral and discount the future at rate λ. There are two phases of life: young and old. Every period young (old) individuals age (die) with probability ζ. All individuals who die are replaced by young unemployed workers. We assume that these shocks are perfectly correlated within a household (father–son pair). This is equivalent to assuming that individuals stop being connected to their parents when they have children, so that at each point in time only two generations are connected.[33]

Workers are indexed by i and differ along several dimensions: preferences for occupations $\boldsymbol{\phi_t^i}$, comparative advantage τ^i, occupation–specific human capital $h^i_{o,t}$ and social capital $n^i_{o,t}$ (networks). Preferences $\boldsymbol{\phi}_t^i$ are represented by a vector of size O, where the o^{th} element is the level of non–pecuniary benefits associated with occupation o. We assume that preferences have both a permanent and a transitory component, so that the period t non–pecuniary benefits are equal to the sum of the two components: $\phi^i_{o,t} = \phi^{i,P}_o + \phi^{i,T}_{o,t}$, for each o. The permanent preferences component $\boldsymbol{\phi}_t^{i,P}$, as well as the comparative advantage τ^i, are drawn at birth (i.e. on entry into the labor market) and do not change over time.[34] In contrast, the temporary component of preferences, as well as occupation–specific human and social

[31] This could also be interpreted as the effect of social pressure or, more generally, any other factor that shifts the utility level for a specific occupation.

[32] This is important, in order to provide the correct mapping between the model and the data (which has an inherently dynamic nature).

[33] This assumption is made for simplicity and does not alter the results.

[34] We do not allow for any investment in types, even though we do ackwnowledge that this may be important for the quantification of mismatch. Doing this in a credible way would require us to introduce assets and borrowing constraints, and to think seriously about occupation heterogeneity and its interaction with educational choice. We leave this to future research.

capital, evolve over time according to laws of motion to be specified below. Each worker is either employed or unemployed ($e_t^i \in \{0,1\}$), and is attached to some occupation o^i.[35] Unemployed workers receive an unemployment benefit equal to b per period.

We denote the father's variables using an F superscript, so that the occupation of individual i's father will be denoted by $o^{i,F}$, the father's networks by $n_o^{i,F}$ and so on.

Upon matching, the surplus generated is split according to a linear sharing rule, such that the wage is set to a share χ of the worker's productivity.[36]

4.2 Search and Relocation across Occupations

We assume that search is costless and directed across occupations. Workers start their working life unemployed and decide in which occupation to look for vacancies. Employed parents help their unemployed sons find a job, by letting them use part of their occupation–specific network. As a consequence, unemployed sons find vacancies in their father's occupation with higher probability than anywhere else, *ceteris paribus*. We assume that unemployed fathers do not help their sons, since they are actively searching for a job themselves.

Each occupation is a separate labor market, where the number of matches between unemployed workers and vacancies is governed by the following constant returns to scale technology:

$$M_{o,t} = (U_{o,t})^{\eta}(V_{o,t})^{1-\eta}, \tag{33}$$

where $M_{o,t}$ denotes the total number of matches produced, $U_{o,t}$ is the total efficiency units of search exerted and $V_{o,t}$ is the measure of vacancies posted at time t in occupation o. The elasticity of the matching function with respect to $U_{o,t}$ is given by η.

Search effort is exerted both by unemployed workers and employed fathers whose sons are currently unemployed. When searching for a job, workers exploit their social networks. Networks are assumed to operate such that information on vacancies can flow within athem at zero cost and there is no competition among workers belonging to the same network. Thus, social networks can help workers find a job, and having a larger network represents an advantage for unemployed workers. This is represented in the model by an increase in the efficiency units of search that a worker can exert. In particular, it is assumed that a worker

[35] Another way of modelling this would be to have the unemployed pool out of all occupations. We claim that this alternative model would yield exactly the same implications as our model, due to the CRS matching function and the fact that we focus on a symmetric equilibrium.

[36] Assuming other forms of wage determinations, such as Generalized Nash Bargaining, would be a rather extreme assumption in our setup, since in that case we would have to assume that all networks, preferences and the productivity levels in *all other* occupations are common knowledge within the match.

with network $n^i_{o,t}$ can exert $(1+n^i_{o,t})$ efficiency units of search. Thus,

$$U_{o,t} = \underbrace{\int \left(1+n^i_{o,t}\right) \mathbb{1}\{e^i = 0; o^i = o\} di}_{\text{Units of Search of Unemployed workers}} + \underbrace{\int \xi \left(1+n^{i,F}_{o,t}\right) \mathbb{1}\{e^i = 0; o^i = o^{i,F} = o; e^{i,F} = 1\} di}_{\text{Units of Search provided by Employed Fathers}}, \tag{34}$$

where ξ represents the proportion of the father's network passed on to the son. The occupation–specific individual job–finding rate $p^i_{o,t}$ is the sum of the probability of an individual finding a job himself (either through his own effort or through his social network) and the probability of his father finding a vacancy for him, if they are in the same occupation:

$$p^i_{o,t} = \frac{M_{o,t}}{U_{o,t}} \left[(1+n^i_{o,t}) + \xi(1+n^{i,F}_{o,t})\mathbb{1}\left\{e^{i,F} = 1; o^{i,F} = o\right\}\right]. \tag{35}$$

Thus, at each time t workers face a single job–finding rate in their current occupation.

We assume free entry of firms, and that posting a vacancy costs κ per period. Firms in occupation o meet with a worker with probability $q_{o,t} = \frac{M_{o,t}}{V_{o,t}}$. Matches are exogenously destroyed at rate δ per period.

Workers (both employed and unemployed) can freely relocate across occupations.[37] It is assumed that when an employed worker decides to relocate, he is separated from his current match (i.e. the match is destroyed) and moves into the unemployment pool of his new occupation. Furthermore, his occupation–specific human capital and social contacts stocks fully depreciate upon changing occupation.

4.3 Intergenerational Transmission and Laws of Motion

We assume that upon entry into the labor market, an individual imperfectly inherits the duplet $\{\boldsymbol{\phi^P}, \tau\}$ from his father:

$$\boldsymbol{\phi^P} \sim F\left(\boldsymbol{\phi^P}|\boldsymbol{\phi^{F,P}}\right) \tag{36}$$

$$\tau \sim G\left(\tau|\tau^F\right) \tag{37}$$

The initial level of occupation–specific human and social capital is assumed to be zero in all occupations. Both human and social capital evolve over time according to the following

[37] We abstract from direct costs of relocation, since these cannot be separately identified from the magnitude of the standard deviation of preference shocks.

laws of motion:

$$h_{o,t+1} = F_h\left(h_{o,t}|e_t\right) \tag{38}$$

$$n_{o,t+1} = F_n\left(n_{o,t}|e_t\right) \tag{39}$$

Finally, the temporary preference vector is drawn each period from the distribution F_ϕ:

$$\boldsymbol{\phi^T}_t \sim F_\phi \tag{40}$$

As already mentioned, we assume that h_o and n_o are reset to zero following a change in occupation.[38]

4.4 Timing

The timing of the model is as follows:

1. Old (young) workers die (age) with probability ζ. A young worker who has aged loses the connection to his father and gives birth to an unemployed son.

2. Preference shocks are realized.

3. Unemployed and employed workers decide whether or not to relocate.

4. Wages and unemployment benefits are paid, and occupation–specific utility flow is realized.

5. Exogenous separations take place. Unemployed workers either find a job or remain unemployed.

6. The workers' state variables are updated according to the laws of motion.

4.5 The Worker's Problem

At the beginning of a worker's life, the worker's problem consists of choosing the occupation in which to search. Besides this initial choice, workers have the option of relocating into a different occupation at the beginning of each period. In what follows, we suppress the i superscript and the t subscript for readability, although all variables (except for $\boldsymbol{\phi^P}$ and τ) change over time. We denote the next period's state variables with a prime. All functional equations are conditional on the worker's state variables.

[38] We do this for computational reasons, even though in principle it would be interesting to track all occupation–specific variables and have them decay over time when the worker is no longer attached to that occupation.

Denote the state of a worker by $\Gamma = \{\boldsymbol{\phi}, \tau, h_o, n_o, o, e\}$, where for simplicity o is set equal to zero for those workers who are choosing an occupation for the first time. A young worker's choices are influenced by his father who can help him find a job, so that his own state also includes his father's state $\Gamma^F = \{\boldsymbol{\phi^F}, \tau^F, h_o^F, n_o^F, o^F, e^F\}$. We must track all of the father's state variables because the son takes into account that: i) even if his father is unemployed today (and therefore does not affect the current job–finding rate), his father will help them him find a job in the future once he becomes employed; and ii) fathers also change occupations over the life–cycle. Conversely, we make the father's problem independent of his son's; that is, a father optimizes his choices without taking into account the impact they have on his son's problem.[39] In the following, we make explicit the dependence of a worker's value functions on his employment status and occupation. Hence, conditional on employment status and occupational affiliation, we denote the state variable of workers by $\Omega = \{\boldsymbol{\phi}, \tau, h_o, n_o\} \cup \Gamma^F$. All Bellman equations are conditional on Ξ, the aggregate state variables, even though we omit this dependance for readability. We first write the value functions for old workers (denoted by a subscript F), with the understanding that they are characterized by $\Gamma^F = \varnothing$.

The Father's Problem

We denote by W^R the value of relocation across occupations:

$$W_F^R(\Omega) = \max_{j \in \{1,\ldots,O\}} \Big\{ W_{j,F}^U(\Omega) + \phi_{j,F}^T \Big\}, \tag{41}$$

where $\phi_{j,F}^T$ represents the temporary preference shock for occupation j. Note that unemployed workers draw a vector of size O–by–1 of preference shocks each period.

The value of unemployment in occupation o (net of the preference shock), $W_{o,F}^U$, includes the value of unemployment benefits for the current period and the expected discounted value of the future:[40]

$$W_{o,F}^U(\Omega) = b + \tilde{\lambda}\Big[p_o(\Omega)\, \mathbb{E}\left[W_{o,F}^E(\Omega')\right] + (1 - p_o(\Omega))\, \mathbb{E}\left[W_{o,F}^R(\Omega')\right]\Big]. \tag{42}$$

[39] From the model's standpoint, this is akin to assuming that fathers are not altruistic (i.e., they attach zero weight to their son's value function). We make this assumption for two reasons: first, we believe that this represents more faithfully actual occupational choices (due to the timing of fertility vs. occupational choices – that is, occupational choices are typically made first). Second, this is unlikely to have a large quantitative effect, given that in the data more than 80% of workers are occupational movers; therefore, fathers expect theat their choices would matter litte for their offspring, in expectation. Third, we do this for simplicity, since allowing for an altruistic motive of fathers would create a complex dynamic game between fathers and sons (for instance, see Barczyk & Kredler 2014).

[40] In this case, the value function for unemployment has to be interpreted at the stage immediately *after* the relocation decision. That is, the worker has to spend the whole period unemployed in occupation o.

An unemployed worker is matched with a vacancy in his occupation with probability $p_o(\Omega)$, and remains unemployed with probability $(1 - p_o(\Omega))$, in which case he can decide to relocate in the next period. The future is discounted at the rate $\tilde{\lambda} = \lambda(1 - \zeta)$, in order to account for the risk of dying.

Employed workers face the relocation decision at the beginning of each period. If they decide to stay on the job, they receive the flow utility, earn the corresponding wage and stay in the same job next period, unless their match is exogenously destroyed (which happens with probability δ). Define $\hat{W}^E_o(\Omega)$ to be the value of staying employed in occupation o (that is, the value of being employed and choosing not to relocate):

$$\hat{W}^E_{o,F}(\Omega) = \phi^P_{o,F} + \phi^{T,E}_{o,F} + w(\Omega) + \tilde{\lambda}\Big[(1-\delta)\,\mathbb{E}\left[W^E_{o,F}(\Omega')\right] + \delta\,\mathbb{E}\left[W^R_F(\Omega')\right]\Big]. \tag{43}$$

At the start of each period, a worker's value function is as follows:

$$W^E_{o,F}(\Omega) = \max\Big\{\hat{W}^E_{o,F}(\Omega), W^R_F(\Omega)\Big\}, \tag{44}$$

since this includes the possibility of leaving the job and relocating into a different occupation.

Notice that employed workers draw two sequences of preference shocks: the first determines whether or not they stay on the job, while the second determines their new occupation, in the case they wish to relocate.[41]

4.6 Wage Determination

Upon matching, wages are set by a linear sharing rule, such that the worker is paid a fixed share of his productivity. Denote $y(\tau, h)$ as the productivity level of a worker of type (τ, h). Then, the equilibrium wage is:

$$w(\tau, h) = \chi\ y(\tau, h).$$

Importantly, we assume that all payoff–relevant information is common knowledge within the match. We also assume that wages are perfectly flexible and that they are renegotiated every period, upon changes in the worker's level of human capital.

[41] This is done for computational convenience.

4.7 Relocation Across Occupations

We denote by j^* the preferred occupation in which to search, namely the occupation that maximizes the value of relocation:

$$j^*_F(\Omega) \in \text{argmax}_{j\in\{1,\dots,O\}}\Big\{W^U_{j,F}(\Omega) + \phi^T_{j,F}\Big\}. \tag{45}$$

Notice that $j^*(\Omega)$ may or may not be the same as a worker's current occupation. If not, then an unemployed worker will always want to relocate, while in the case of an employed worker the choice will depend on the difference between the value functions $\hat{W}^E_{o,F}(\Omega)$ and $W^R_F(\Omega)$.

We define $R^k_{o,F}(\Omega)$ (for $k \in \{E, U\}$) as the policy function with respect to the relocation decision. Thus, when $R^k_{o,F}(\Omega) = 1$, a worker of type Ω with employment status k in occupation o optimally decides to relocate. In other words:

$$R^U_{o,F}(\Omega) = \mathbb{1}\{j^*_F(\Omega) \neq o\}.$$

$$R^E_{o,F}(\Omega) = \mathbb{1}\{W^R_F(\Omega) > \hat{W}^E_{o,F}(\Omega)\}.$$

The Son's Problem

A son faces a very similar problem to that of a father. The only difference is that he takes into account his father's decisions. As a result, a young worker can decide to relocate as a consequence of a change in his own state variables (preferences) or because his father's state variables have changed, in which case he might want to follow his father in order to benefit from a higher probability of finding a job.

The expression for the value of relocation remains identical:

$$W^R_S(\Omega) = \max_{j\in\{1,\dots,O\}} \Big\{W^U_{j,S}(\Omega) + \phi^T_{j,S}\Big\}. \tag{46}$$

The value of unemployment and employment are augmented by the fact that the worker will become a father in the next period with probability ζ:

$$\begin{aligned} W^U_{o,S}(\Omega) = b + \lambda\Big[&p_o(\Omega)\ \big(\zeta\,\mathbb{E}\left[W^E_{o,F}(\Omega')\right] + (1-\zeta)\mathbb{E}\left[W^E_{o,S}(\Omega')\right]\big) \\ &+(1-p_o(\Omega))\ \big(\zeta\,\mathbb{E}\left[W^R_{o,F}(\Omega')\right] + (1-\zeta)\mathbb{E}\left[W^R_{o,S}(\Omega')\right]\big)\Big]. \end{aligned} \tag{47}$$

$$
\begin{aligned}
W^E_{o,S}(\Omega) = \max \Big\{ & \phi^P_{o,S} + \phi^{T,E}_{o,S} + w(\Omega) \\
& + \lambda \Big[(1-\delta) \left(\zeta \, \mathbb{E}\left[W^E_{o,F}(\Omega')\right] + (1-\zeta)\mathbb{E}\left[W^E_{o,S}(\Omega')\right] \right) \\
& + \delta \left(\zeta \, \mathbb{E}\left[W^R_F(\Omega')\right] + (1-\zeta)\mathbb{E}\left[W^R_S(\Omega')\right] \right) \Big], W^R_S(\Omega) \Big\}.
\end{aligned} \tag{48}
$$

The relocation decisions $R^U_{o,S}(\Omega)$ and $R^E_{o,S}(\Omega)$ are isomorphic to those of the father, and are defined according to the above–specified value functions.

4.8 The Firm's Problem

A firm is represented by a single job that is either filled or vacant. The value function for a job filled with a worker of type Ω is denoted by $J_{o,k}(\Omega)$, where $k \in \{F, S\}$ denotes the age of the worker. Provided that the worker does not choose to leave the firm, this value function includes the current profit (given by production net of the wage payment) and the continuation value of keeping the worker. The value of keeping an old worker is given by:

$$
\begin{aligned}
J_{o,F}(\Omega) = & \, \mathbb{1}\{R^E_{o,F}(\Omega) = 0\} \Big[y(\tau, h_o) - w(\Omega) + \tilde{\lambda} \big[(1-\delta)\mathbb{E}\left[J_{o,F}(\Omega')\right] + \delta V'_o \big] \Big] \\
& + \mathbb{1}\{R^E_{o,F}(\Omega) = 1\} \, V_o.
\end{aligned} \tag{49}
$$

The output of the match is given by the function $y(\tau, h_o)$, which we assume to be increasing in both arguments. The match is exogenously destroyed with probability δ in the next period (in which case, as in the case of endogenous separation, the firm is left with the value of a vacancy V_o). With probability $(1-\delta)$, the match continues, and the state variables of the worker are updated.

The value of keeping a young worker is as follows:

$$
\begin{aligned}
J_{o,S}(\Omega) = & \, \mathbb{1}\{R^E_{o,S}(\Omega) = 0\} \Big[y(\tau, h_o) - w(\Omega) + \lambda \big[(1-\delta) \big(\zeta \, \mathbb{E}\left[J_{o,F}(\Omega')\right] \\
& + (1-\zeta)\mathbb{E}\left[J_{o,S}(\Omega')\right] \big) + \delta V'_o \big] \Big] + \mathbb{1}\{R^E_{o,S}(\Omega) = 1\} \, V_o.
\end{aligned} \tag{50}
$$

This equation has the same interpretation as the one for an old worker, except that it allows for the possibility of a worker becoming old (as a result of the ζ shock) and the match continuing.

The value of a vacancy V_o is given by the expected profits less the posting cost κ.

$$V_o = -\kappa + q_o \mathbb{E}\left[J_o(\Omega')\right] + (1-q_o)V_o', \tag{51}$$

where the expectation is taken over the distribution of unemployed workers in occupation o, which include all possible types Ω and possible ages $\{F, S\}$.

4.9 Equilibrium Definition

We focus on a steady state equilibrium (SS) in which all value functions and relocation decisions are constant over time. As a result, worker flows are also constant over time (the equations describing such flows are relegated to the Appendix).

Definition: An SS equilibrium is a set of value functions $W^U_{o,F}(\Omega)$, $W^U_{o,S}(\Omega)$, $W^E_{o,F}(\Omega)$, $W^E_{o,S}(\Omega)$, V_o; relocation decisions $R^U_{o,F}(\Omega)$, $R^U_{o,S}(\Omega)$, $R^E_{o,F}(\Omega)$, $R^E_{o,S}(\Omega)$, $j^*_F(\Omega)$, $j^*_S(\Omega)$; labor market tightness θ_o; wages $w_o(\Omega)$; laws of motion for the individual state variables; and laws of motion of unemployed and employed workers over all occupations, such that:

- The value functions for workers and relocation decisions satisfy Equations (42), (44), (47), (48).
- There is free entry into all occupations: $V_o = 0 \ \forall o \in \{1, .., O\}$.
- Labor market tightness satisfies Equation (51).
- Wages satisfy Equation (31).
- Individual state variables evolve according to Equations (38), (39) and (40).
- Distributions of workers evolve according to Equations (63), (64), (65) and (66) (in the Appendix).
- The measures of employed and unemployed workers of each type Ω are constant over time.
- Flows of employed and unemployed workers of each type Ω are constant over time.

5 Quantitative Analysis

In this section, we quantitatively assess the importance of each of the channels operating in the model (ability, preferences and networks transmission) in delivering occupational persistence. We first assign values to the structural parameters of our model, and then use the calibrated model to decompose occupational persistence and perform welfare analysis and policy experiments.

5.1 Calibration Strategy

Our strategy involves exogenously fixing some of the parameters, and jointly calibrating all the rest to relevant moments of the UK data. First, we fix the number of occupations O to 9 in order to be consistent with the SOC–1 digit aggregation. Each occupation attracts in equilibrium the same amount of workers, with an identical composition of productivity, preferences and networks. This arises because the productivity and preferences distributions, the matching functions and the vacancy posting costs are all symmetric across occupations. One period in the model corresponds to one month, and therefore the discount factor λ is set to 0.9966. The age shock ζ is set so as to deliver an average working life of 40 years (20 as a young worker and 20 as an old one), implying a value of 0.00416. We also fix the surplus sharing rule parameter χ to 0.7 and the scale of the matching function A to 0.1. Finally, we fix $\eta = 0.5$ following Petrongolo and Pissarides (2001).

We calibrate the rest of the parameters in order to match relevant features of the data. In order to do so, we first need to choose the grid of possible values of the worker–specific state variables, as well as the functional forms describing their laws of motion. We let h and n take two different values, with the lower one being normalized to 1:

$$h \in \{1, 1 + \hat{h}\},$$

$$n \in \{1, 1 + \hat{n}\},$$

where $\hat{h}$ and $\hat{n}$ represent the premia of human capital and networks that are associated with tenure.[42] The accumulation/depreciation of these occupation–specific variables is subject to a Markov-process characterized by the following parameters: $p_h^+, p_h^-, p_n^+, p_n^-$, where the $+$ and $-$ superscripts denote accumulation (when employed) and depreciation (when unemployed) probabilities, respectively.

We also assume that each worker has a talent τ in one occupation, in which he has a productivity premium $\hat{\tau}$. The minimum level of productivity is normalized to 1. The match

[42] This choice allows us to avoid irrealistically high values of accumulated variables, that would otherwise arise due to the stochastic ageing assumption.

production function $y(\tau, h_o)$ is assumed to be:

$$y(\tau, h_o) = \begin{cases} h_o & \text{if} \quad o \neq \tau \\ \hat{\tau}\, h_o & \text{if} \quad o = \tau \end{cases}$$

In the same way, we assume that each worker has a preferred occupation ϕ, where he obtains a non–pecuniary benefit ($\hat{\phi}$) that is higher than elsewhere. At the same time, we also normalize the baseline level of preferences for a job $\bar{\phi}$ to 0. Both τ and ϕ are drawn at birth, with ρ_τ and ρ_ϕ being the probabilities of drawing the same values as the father ($\rho = 1$ represents perfect persistence). Let τ and ϕ be respectively the occupations in which a worker has a comparative advantage and the preference premium:

$$\tau = \begin{cases} \tau^F & \text{w.p.} \quad \rho_\tau \\ o \neq \tau^F & \text{w.p.} \quad \frac{(1-\rho_\tau)}{O} \quad \forall\, o \neq \tau^F \end{cases}$$

$$\phi = \begin{cases} \phi^F & \text{w.p.} \quad \rho_\phi \\ o \neq \phi^F & \text{w.p.} \quad \frac{(1-\rho_\phi)}{O} \quad \forall\, o \neq \phi^F \end{cases}$$

Finally, we assume that the idiosyncratic preference shocks are drawn from a Type–1 Extreme distribution, with standard deviation σ.[43]

Together with κ and ξ, we have a total of 14 parameters to be calibrated. We search for the parameter configuration that minimizes the following loss function:

$$\mathcal{L} = \frac{\sqrt{\sum_{n=1}^{K} \left(\frac{M_n(\Theta) - T_n}{T_n}\right)^2}}{K},$$

where T is a K–by–1 vector containing our target statistics and M is a K–by–1 vector containing the statistics generated by the model (we choose $K = 14$, so that the model is exactly identified). Table VI contains the list of all parameters of the model, each of which is associated with the identifying moment in Column 4.

The vacancy posting cost κ is calibrated in order to match the average monthly UE rate, which is 0.1251. A lower posting cost induces more entry from the firms' side, implying higher tightness and higher finding rates. The exogenous separation rate δ is set in order to match the average EU rate, which is 0.0047. The transmission of networks ξ is set to replicate the job–finding rate premium (of 0.0546) of occupational followers w.r.t. movers. A higher value of ξ implies that a son can take advantage of a larger proportion of his

[43] This is a standard assumption in the literature on occupational choice. See for instance Wiczer (2014).

father's network. The transmission of comparative advantage is calibrated to match the differential in the likelihood of occupational persistence by the father's wage. An increase in ρ_τ increases the chances that the occupation of the father is also that in which the son finds his comparative advantage *when the father is well–matched*, thus increasing persistence for those with a high–wage father.[44] Therefore, we target the difference in probability of being an occupational follower if the father's wage is above the average, as observed in the BHPS data. To obtain this target, we regress $\pi_{i,t}$ on a dummy taking value 1 if the father's log wage is above the average and zero otherwise, controlling for our usual covariates.[45] We find that there is a 2.3% difference in the probability of being a follower between high-wage and low-wage fathers (details of the estimation are reported in Table XX in the Appendix).[46] The parameter governing the transmission of preferences ρ_ϕ is pinned down by asking the model to replicate the occupational persistence observed in the data (likelihood ratio of 1.72 at the 1–digit level). In other words, we are using the transmission of preferences as the residual channel to entirely match occupational persistence, above and beyond the persistence already generated by the other two channels.

Next, the comparative advantage premium $\hat{\tau}$ is calibrated to match the level of within–occupation log wage variance. The rationale for this is that the more heterogeneous are the potential productivity levels of workers across occupations, the more dispersed equilibrium wages will be. The networks premium $\hat{n}$ is calibrated to match the proportion of jobs found through networks in the UK, which is 0.23 (Pellizzari 2010). The higher $\hat{n}$ is, the more networks will be present in the economy and used for job search. The preference premium $\hat{\phi}$ is chosen to replicate the average wage discount (of 0.0763 log points) of occupational followers. High values of $\hat{\phi}$ imply that preferences are relatively more important than comparative advantages in occupational choice. The standard deviation of the preference shocks distribution (σ) is calibrated to the probability of switching occupation after an unemployment spell (0.3567). The larger the variance of the shocks, the more frequently occupational changes occur. The value of unemployment b is calibrated to match the average replacement rate in the UK of 0.53 (OECD).

Moreover, we calibrate $p_h^+ = 0.0166$ which, together with $\hat{h} = 0.26$, implies that the average occupational returns after 5 years are equal to 26%, as observed in the data. The probability of losing human capital p_h^- is calibrated to match the average wage discount after

[44]This was already explained, more in length, in Section 3.7.

[45]We do not use the estimates of Table XIX, which were made treating the father's wage as a continuous variable, only because our model features only four wage levels. Therefore, we believe that looking at fathers with relatively high and relatively low wage levels provides a better mapping to the data.

[46]We also run the same estimation dividing fathers in those above their occupation-specific average log wage and those below, and our results are substantially unchanged (Column 2 of Table Table XX in the Appendix).

unemployment of 7.6% (Arulampalam 2001). The probability of losing networks p_n^- is set to match the slope of the JF rate–unemployment duration profile. In particular, we ask the model to replicate the drop in finding rates that occur between the first and second months of unemployment duration. Finally, we calibrate the probability of accumulating networks p_n^+ to the conditional correlation of job–finding rates with months of past occupational tenure, which is 0.008.

5.2 Calibration Results

The model is able to precisely match all targets, both the cross–sectional and the inter–generational ones. We are able to replicate the full extent of occupational persistence observed in the data by making both preferences and comparative advantages persistent across generations (where the probability of inheriting them is 0.147 and 0.141, respectively).The proportion of parental networks exploited by the son is 0.325, which generates the same job–finding rate premium as in the data.

A large degree of heterogeneity is needed in order to match the data moments: the preference premium is 0.811, while the comparative advantage premium is even higher, at 1.008. The networks premium is also substantial (1.104), whereas the human capital premium is 0.26 (taken directly from the data). The accumulation of both human capital and networks is slow: the monthly probability of human capital growing is 0.0166, while for networks it is 0.005. In contrast, their depreciation during unemployment is substantially faster: the monthly probability of networks depreciating is 0.115, while for human capital it is 0.79.

We calculate that in this economy posting a vacancy costs around 4 times the average wage. Finally, the exogenous match destruction rate is 0.003, with the rest of the EU flows being accounted for by endogenous separations.

Parameter	Description	Value	Target/Source	Data	Model
Intergenerational Transmission					
ξ	Transmission of Networks	0.325	JF premium of followers (BHPS)	0.055	0.055
ρ_τ	Transmission of Comparative Advantage	0.141	Difference in Proportion of Followers by Father's wage (BHPS)	0.023	0.023
ρ_ϕ	Transmission of Preferences	0.147	Interg. Occupational Persistence (Lik. Ratio, BHPS)	1.720	1.720
Heterogeneity and Laws of Motions					
$\hat{h}$	HC premium	0.260	Average occupational tenure returns after 5 years (BHPS)	-	-
$\hat{n}$	Networks premium	1.104	Proportion of jobs found through contacts (Pellizzari 2010)	0.230	0.230
$\hat{\tau}$	Comparative Advantage premium	1.008	Within–occupation log wage variance (BHPS)	0.166	0.166
$\bar{\phi}$	Baseline preference for jobs	0	Normalization	-	-
$\hat{\phi}$	Preference premium	0.811	Wage discount of followers (BHPS)	0.076	0.076
p_h^+	Probability of accumulating HC (employed)	0.0166	Average occupational tenure returns after 5 years (BHPS)	0.260	0.260
p_h^-	Probability of losing HC (unemployed)	0.790	Average wage discount after unemp. (Arulampalam 2001)	0.076	0.076
p_n^+	Probability of accumulating networks (employed)	0.005	Regression of JF rate vs. past occupational tenure (BHPS)	0.008	0.008
p_n^-	Probability of losing networks (unemployed)	0.115	JF rate–unemployment duration profile (see text, BHPS)	1.066	1.066
σ	Standard deviation of preference shocks	0.254	Occupational change rate, after U spell (monthly, BHPS)	0.357	0.357
Environment					
O	Number of occupations	9	1-digit SOC aggregation	-	-
κ	Vacancy posting cost	5.269	Average UE rate (monthly, BHPS)	0.125	0.125
δ	Exogenous separation rate	0.003	Average EU rate (monthly, BHPS)	0.005	0.005
λ	Discount factor	0.9966	From literature	-	-
ζ	Age shock	0.00416	Average length of worklife: 20 (young) + 20 (old) years	40	40
b	Unemployment benefit	0.745	Average replacement rate (OECD)	0.530	0.530
χ	Surplus sharing rule	0.7	Normalization	-	-
A	TFP parameter of matching function	0.1	Normalization	-	-
γ	Elasticity of matching function w.r.t. unemp.	0.5	Petrongolo and Pissarides (2001)	-	-

Table VI. Calibration Results

5.3 Occupational Persistence Decomposition and Welfare Analysis

The model allows us to study the factors behind occupational choice, and how they differ in importance between followers and movers. In Table VII, we calculate how often the occupational choice is aligned with each of the three possible factors (parental networks, comparative advantage, and preferences) under the baseline calibration.

	All	**Followers**	**Movers**
Sorting along comparative advantage (fathers)	0.656	-	-
Sorting along preferences (fathers)	0.455	-	-
Sorting along parental networks (sons)	0.183	0.959	0.000
Sorting along comparative advantage (sons)	0.708	0.624	0.729
Sorting along preferences (sons)	0.402	0.461	0.388
Average log wage (sons)	0.292	0.232	0.307
Average unemployment rate (sons)	0.061	0.045	0.065

Table VII. Occupational sorting.

For fathers, comparative advantage seems to be more important than preferences for occupational sorting: 66% (46%) of fathers choose the occupation in which they have a comparative advantage (preference). Among sons, the same holds true: about 71% of them pick the occupation in which they are most productive, whereas about 40% of them pick their preferred occupation. Finally, the occupational choice is aligned with parental networks in 18% of the cases. Striking differences in sorting arise between followers and movers: the former put more weight on preferences in their occupational decision (46% versus 39% of movers) and less on comparative advantage (62% versus 73% of movers). As a consequence, followers earn lower wages, as can be seen in row 6 of Table VII. At the same time, followers have better employment prospects than movers, with an average unemployment rate of 4.5%, versus 6.5% for movers. Summing up, the model economy generates a clear sorting of workers in the two regions of high–employment/low–wages and low–employment/high–wages.

However suggestive, these correlations are not yet informative about the nature of occupational persistence. For this reason, we now sequentially shut down each of the three channels delivering occupational persistence. In this way, we are able to: i) quantify the contribution of each channel to overall persistence; and ii) evaluate welfare in each different scenario. To evaluate welfare in Steady State, we use the following function:

$$\mathcal{W} = O \int_{\Omega} \left[(1 - u(\Omega)) \left[y(\Omega) + \phi(\Omega) \right] + u(\Omega) b - \kappa \theta \right] \, dF\Omega, \tag{52}$$

where $u(\Omega)$, $y(\Omega)$ and $\phi(\Omega)$ represent the equilibrium unemployment rate, the productivity level and the preference component of a given type Ω, respectively. Due to the symmetry of the equilibrium, the aggregation across occupations is achieved through a simple multiplication.

Table VIII shows the results of the experiments: column 1 is the case of the baseline economy, while in columns 2-8 we set $\xi = 0$, $\rho_\tau = 1/O$ and $\rho_\phi = 1/O$, along with all possible combinations of these parameter changes. First, all factors seem to matter for occupational persistence, though by differing degrees. Shutting down parental networks generates the largest drop in persistence, of about 79% (column 2), while comparative advantage and preferences transmission (columns 3 and 4) respectively account for about 19% and 10% of persistence. Moreover, networks transmission appears to work in conjunction with the other sources of persistence, since shutting down these channels in pairs delivers less of a drop than the sum of the effects separately (columns 5 and 6 vs 2–4). In contrast, comparative advantage and preferences work independently from one another (the drop in column 7 is equal to or even larger than the combined effects of column 3 and 4). To better understand the surprisingly large effect of networks and how they interact with the other factors, in Figure 4 we plot the average policy function (occupational choice) of unemployed workers whose father is employed and whose comparative advantage and preferences are not aligned.[47]

As already noted, individuals in this economy tend to choose their occupation more according to comparative advantage than preferences. Moreover, the presence of an employed father strongly impacts the occupational choice of his son. For instance, on average, individuals choose the occupation in which they have a comparative advantage with a probability of 80% if the father is also employed in that occupation. This probability drops to 60% if the father is employed in a different occupation (compare the first two bars in Figure 4). This effect is even larger for preferences: the occupation for which preference and parental networks are aligned is chosen in 55% of the cases, while the preferred occupation without parental networks is chosen in only 26% of the cases. It is significant that the benefits from the father's networks alone are not enough to attract the son. Indeed, by comparing the last two bars, one can easily see that choosing an occupation with neither comparative advantage nor preferences is almost never an attractive option, with or without the father's network. The reason for this stark difference is that the value of employment differs from the value

[47] The workers for which comparative advantage and preferences are not aligned represent the large majority of the population. In the Appendix, we show the same average policy function of those workers for whom the two factors are aligned.

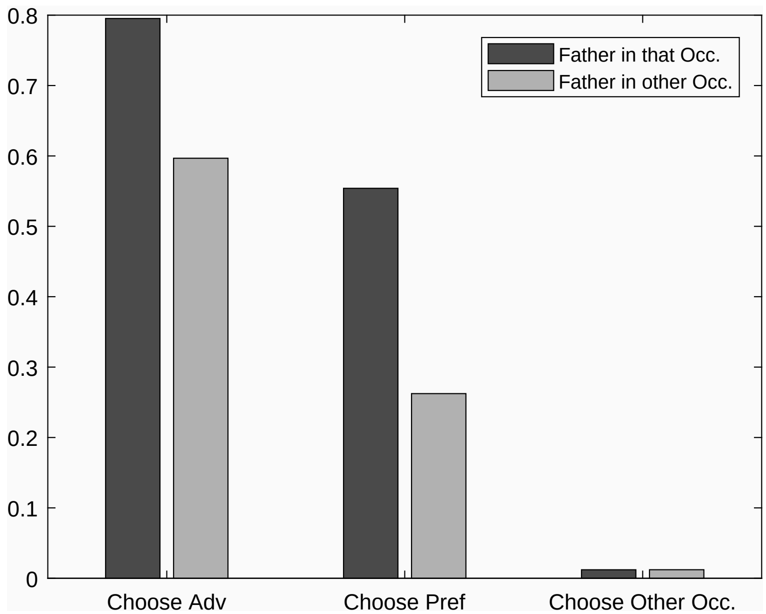

Figure 4. Probability of choosing occupations (Average Policy Function) for unemployed workers with comparative advantage and preference in different occupations.

of unemployment to a larger extent in occupations with either comparative advantage or preference than in other occupations. By improving the chances of employment, parental networks act as a multiplier of these differentials, therefore playing a much larger role in conjunction with these other fixed factors than alone.

Second, the welfare consequences of a reduction in persistence vary widely across the experiments. When we shut down parental networks (column 2), welfare improves by 0.11%, due to the improved allocation of workers to occupations (sorting along the productivity dimension increases from 71% to 74%) and despite a worsened sorting along the preferences dimension (which drops from 40% to 37%). As a consequence of the increase in the productivity of the workforce, output per worker increases and wage variance decreases. Also, unemployment improves (declines by 1.5%) despite the fact that less efficiency units of search are now exerted in the market, since firms react to the change in average labor productivity by posting more vacancies. Overall, the welfare change is small because most of the improvement along the productivity dimension is undone by the worsened sorting along the preferences dimension. In contrast, when we shut down the transmission of comparative advantage (column 2), welfare decreases by 0.04%, while output per worker declines (sorting along the productivity dimension worsens, while sorting along the preferences dimension improves) and unemployment rises (by 0.33%). Finally, shutting down preferences transmission (column 3) has a similar though smaller effect to that of shutting down parental networks. Thus, productivity becomes more dominant in an individual's choice, output per worker increases and unemployment decreases. The net effect of these changes, despite a worsened sorting along the preferences dimension, is an increase in welfare of 0.03%.

5.4 Other Counterfactual Experiments

5.4.1 The Importance of Multiple Transmission Channels

The model is characterized by several degreess of heterogeneity, and intergenerational persistence is influenced by three different factors (comparative advantage, preferences, networks). One interesting exercise we will now carry out is to shut down some of these channels and recalibrate the model in order to match the data with fewer degrees of freedom. This allows us to understand whether all model dimensions are really necessary in order to replicate the data patterns. We keep the transmission of productive abilities as the only transmission channel, since it embeds in a reduced–form way genetic transmission, educational choices and human capital transfers in general, which are the channels most commonly emphasized in the intergenerational literature. Therefore, we set $\xi = 0$ and $\rho_\phi = 1/O$ and ask the model to match all data moments in Table VI except for the JF rate premium and the wage discount. The rationale for our choice is that, with only one source of persistence, the model cannot replicate either of these two moments.

In general, the fit of the model is now substantially worse.[48] The model is not able to fully account for occupational persistence, producing a likelihood ratio of only 1.597. The value of ρ_τ is set as high as possible, since this is now the only source of occupational persistence. By doing so, the model largely overshoots the differentials in the propensity to be a follower by the father's wage (in fact, it is more than 10 times larger than in the data). Another consequence is that, in order to generate high persistence, the model completely fails to generate the wage discount (non–targeted) of followers relative to movers (and actually generates a wage premium). This reflects the fact that productivity transmission is the only channel producing persistence, and therefore occupational followers base their occupational choice on productivity to a larger extent than movers. By construction, the model also cannot replicate the job–finding rate premium of followers (non–targeted), since networks transmission is shut down.

When we shut down the persistence of comparative advantage in the restricted economy (Table IX), it turns out that persistence is absolutely neutral in this economy. Shutting down the only source of persistence delivers an identical economy in all dimensions, except for occupational persistence, which vanishes completely. This is because in this economy persistence is not a sign of distortions in the occupational choice of individuals. Furthermore,

[48]The calibration table of the restricted model can be found in Appendix C.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
No parental net.	($\xi = 0$)	-	✓	-	-	✓	✓	-	✓
No comp. adv. trans.	($\rho_\tau = 1/O$)	-	-	✓	-	✓	-	✓	✓
No pref. trans.	($\rho_\phi = 1/O$)	-	-	-	✓	-	✓	✓	✓
Occupational Persistence		1.720	1.154	1.582	1.649	1.035	1.119	1.508	1.000
(Δ% from baseline)		0.000	(-78.60)	(-19.23)	(-9.84)	(-95.10)	(-83.50)	(-29.37)	(-100.00)
Welfare (Δ% from baseline)		0.000	(0.112)	(-0.044)	(0.034)	(0.112)	(0.112)	(-0.010)	(0.112)
Sorting along comparative advantage (sons)		0.708	0.737	0.704	0.714	0.737	0.737	0.709	0.737
Sorting along preferences (sons)		0.402	0.373	0.407	0.397	0.373	0.374	0.401	0.373
Sorting along comparative advantage (fathers)		0.656	0.674	0.653	0.660	0.674	0.674	0.657	0.674
Sorting along preferences (fathers)		0.455	0.436	0.458	0.451	0.437	0.437	0.454	0.437
Output per worker (=1 in baseline)		1.000	1.017	0.997	1.003	1.017	1.017	1.001	1.017
Log Wage Variance (Δ% from baseline)		0.166	(-4.230)	(0.640)	(-0.789)	(-4.230)	(-4.230)	(-0.137)	(-4.230)
Welfare CV (Δ% from baseline)		0.147	(-0.111)	(0.022)	(-0.035)	(-0.111)	(-0.111)	(-0.0126)	(-0.111)
Unemployment Rate (Δ% from baseline)		0.061	(-1.505)	(0.334)	(-0.296)	(-1.506)	(-1.506)	(0.035)	(-1.506)
Average UE Rate (Δ% from baseline)		0.125	(1.728)	(-0.289)	(0.314)	(1.728)	(1.728)	(0.028)	(1.728)
Average EU Rate (Δ% from baseline)		0.005	(-0.117)	(0.052)	(0.030)	(-0.118)	(-0.118)	(0.083)	(-0.118)
Equilibrium Tightness		1.565	(3.486)	(-0.577)	(0.628)	(3.486)	(3.486)	(0.057)	(3.486)

Table VIII. Occupational Persistence Decomposition and Welfare Analysis.

persistence in this economy is generated only by the fact that father–son pairs tend to be more similar than two randomly picked workers. In this sense, occupational persistence is no longer a reflection of the fact that sons *care* about the occupational choices of their father and are affected by them. In other words, a son's policy and value functions are now independent of his father's state variables.

		(1)	(2)
No parental net.	($\xi = 0$)	n.a.	n.a.
No comp. adv. trans.	($\rho_{\mathcal{T}} = 1/O$)	-	✓
No pref. trans.	($\rho_{\phi} = 1/O$)	n.a.	n.a.
Occupational Persistence		1.597	1.000
(Δ% from baseline)		-	(-100.00)
Welfare (Δ% from baseline)		-	(0.000)
Sorting along comparative advantage (sons)		0.344	0.344
Sorting along preferences (sons)		0.767	0.767
Sorting along comparative advantage (fathers)		0.365	0.365
Sorting along preferences (fathers)		0.746	0.746
Output per worker (=1 in baseline)		1.000	1.000
Log Wage Variance (Δ% from baseline)		0.181	(0.000)
Welfare CV (Δ% from baseline)		0.153	(-0.000)
Unemployment Rate (Δ% from baseline)		0.067	(-0.000)
Average UE Rate (Δ% from baseline)		0.112	(0.000)
Average EU Rate (Δ% from baseline)		0.004	(-0.000)
Equilibrium Tightness		1.255	(0.000)

Table IX. Occupational Persistence Decomposition and Welfare Analysis (Restricted Model).

5.4.2 The Role of Search Frictions

Search frictions are an important determinant of productive mismatch in our framework. Therefore, it is interesting to investigate the extent to which the severity of frictions affects the importance of parental networks, the level of persistence and the overall allocation. To do so, we impose the degree of frictions implied by the monthly finding rates of different economies on the UK baseline calibration. We focus on two polar cases among OECD countries: the US and Spain. We target the average monthly job–finding rates estimated in Hobijn & Şahin (2009): 0.5630 for the US and 0.0389 for Spain. We recalibrate κ in order to match these rates, keeping all other parameters constant; the implied new values of the parameter are 1.30 (for the US) and 15 (for Spain). We repeat the persistence decomposition exercises of subsection 5.3 for both of the counterfactual economies, with the results shown in Table XXII in the Appendix.

Two main results stand out: First, the importance of parental networks crucially depends on the size of the frictions. In the low–friction economy, removing parental networks barely affects persistence (which is reduced by only 2.3%), whereas the reduction in the high–friction economy is much more pronounced (79%). At the same time, the removal of networks is welfare–improving in the high–friction economy (since it raises average labor productivity), but is welfare–decreasing in the low–friction economy (since it crowds out occupational choice along the preferences dimension, due to the fact that networks are basically not generating any occupational choice that is not based on productivity in the baseline equilibrium). Relatedly, we find that occupational persistence is much higher in the high–friction economy than in the low–friction economy, other things being equal (likelihood ratio of 1.68 vs. 1.25).

Second, by comparing column 1 to column 5, we can see that search frictions may be responsible for high unemployment and low productivity at the same time. This is a reflection of the fact that networks are more distortionary in environments with large frictions, where individuals are more willing to trade their productive advantage for better employment prospects.

5.4.3 Policy Experiment: Unemployment Benefits

We now look at how changes in unemployment benefits affect the equilibrium of the economy. In order to assess the welfare consequences of such changes, we introduce a lump–sum tax τ on existing matches (which is split between workers and firms, where χ is the share paid by the workers) and a government budget constraint. The new value functions for employed workers and firms are as follows:

$$W_o^E(\Omega) = \max\Big\{\phi_o^P + \phi_o^{T,E} + w(\Omega) - \chi\tau + \tilde{\lambda}\Big[(1-\delta)\,\mathbb{E}\left[W_o^E(\Omega')\right] + \delta\,\mathbb{E}\left[W^R(\Omega')\right]\Big], W^R(\Omega)\Big\}. \tag{53}$$

$$J_o(\Omega) = \mathbb{1}\{R_o^E(\Omega) = 0\}\Big[y(\tau, h_o) - w(\Omega) - (1-\chi)\tau + \tilde{\lambda}\big[(1-\delta)\mathbb{E}\left[J_o(\Omega')\right] + \delta V_o'\big]\Big] \quad (54)$$
$$+\mathbb{1}\{R_o^E(\Omega) = 1\}\, V_o.$$

The government balances its budget in each period. That is, the change in unemployment benefits from the baseline equilibrium must be financed by the tax revenues:

$$\Delta b\, u = \tau(1-u), \quad (55)$$

where u is the unemployment rate of the economy. The rest of the model remains unchanged.

Some of the channels commonly emphasized in the literature through which unemployment benefits have an effect on the economy, such as the scope for redistribution (in the presence of risk aversion) or the disincentivizing effect on the search efficiency units choice, are absent in our framework. At the same time, unemployment benefits interact very strongly with the main tradeoff at work in our model. Thus, an increase (decrease) in the value of unemployment benefits decreases (increases) the distance between the value of employment and unemployment for workers. As a consequence, parental networks become less (more) important in the son's choice, since insurance against unemployment becomes less (more) valuable. This implies that workers sort more (less) according to productivity and preferences. To the extent that this increase in sorting is more prominent along the comparative advantage dimension, unemployment benefits can produce productivity gains.[49]

In the quantitative experiment (Table X), an increase of 10% (25%) in b favours sorting along the preferences dimension, whereas it slightly dampens the sorting along the comparative advantage dimension. As a consequence, output per worker decreases by about 0.1% (0.3%). At the same time, occupational persistence decreases (since parental networks are less attractive) and unemployment increases (since unemployment is now a more attractive option). The overall net effect on welfare is negative (−0.09% and −0.31%), reflecting the fact that the tax rate increases proportionally more than b, given that unemployment increases. Columns 4 and 5 show that decrease in b have qualitatively opposite effects.

When we repeat the same exercises under the restricted calibration (lower part of Table X), we find the effect to be similar. One important difference is that, under the restricted calibration, increases in b do indeed lead to improvements in output per worker, even if the magnitude of the change is quite small (0.2% and 0.4%). The increase in unemployment is therefore smaller than in the baseline case, reflecting a relatively higher level of firm entry (in response to the increase in labor productivity). Interestingly, welfare moves in the same direction in both model specifications, even though the magnitude of the change is very

[49] This mechanism has a very similar flavor to that in Acemoglu & Shimer (2000) and Golosov *et al.* (2013).

different. For instance, an increase of 25% in b generates a welfare loss of 0.31% under the baseline calibration, while in the restricted calibration the loss is only 0.16% . Hence, it turns out that allowing for multiple sources of persistence is also relevant for the assessment of labor market policy in general.

	(1)	(2)	(3)	(4)	(5)
Change in b (Δ% from baseline)	-	+10%	+25%	-10%	-25%
Baseline Economy					
Occupational Persistence	1.720	1.702	1.673	1.737	1.762
(Δ% from baseline)	-	(-2.516)	(-6.500)	(2.418)	(5.875)
Welfare (Δ% from baseline)	-	(-0.091)	(-0.311)	(0.058)	(0.094)
Sorting along comparative advantage (sons)	0.708	0.707	0.704	0.709	0.710
Sorting along preferences (sons)	0.402	0.404	0.407	0.402	0.401
Sorting along comparative advantage (fathers)	0.656	0.655	0.652	0.657	0.658
Sorting along preferences (fathers)	0.455	0.456	0.459	0.454	0.454
Output per worker (=1 in baseline)	1.000	0.999	0.997	1.001	1.001
Unemployment Rate (Δ% from baseline)	0.061	(3.694)	(11.369)	(-2.884)	(-6.061)
Restricted Model					
Occupational Persistence	1.597	1.607	1.623	1.589	1.578
(Δ% from baseline)	-	(1.545)	(4.224)	(-1.380)	(-3.185)
Welfare (Δ% from baseline)	-	(-0.054)	(-0.164)	(0.040)	(0.080)
Sorting along comparative advantage (sons)	0.344	0.346	0.350	0.342	0.340
Sorting along preferences (sons)	0.767	0.765	0.761	0.769	0.771
Sorting along comparative advantage (fathers)	0.365	0.367	0.370	0.363	0.361
Sorting along preferences (fathers)	0.746	0.744	0.741	0.748	0.750
Output per worker (=1 in baseline)	1.000	1.002	1.004	0.999	0.997
Unemployment Rate (Δ% from baseline)	0.067	(2.476)	(6.925)	(-2.156)	(-4.890)

Table X. Policy Experiment: effect of changes in unemployment benefits.

6 Conclusions

We investigated the determinants of occupational persistence across generations. When persistence is generated from multiple sources, it is crucial to assess their relative importance in order to understand the relationship between persistence and misallocation, and to derive welfare implications. A simple model of occupational persistence and search frictions, in which both abilities and contacts are transmitted across generations, delivers clear–cut testable predictions on employment prospects and wages, which are confirmed in UK data.

We extended the theory to a more complete dynamic model of occupational choice, allowing for mobility over the life–cycle and accumulation/depreciation of human and social capital. We found that parental networks account for the bulk of occupational persistence and that a model based only on transmission of ability (the restricted model) would be at odds with several features of the data. A key result of our quantitative analysis is that only occupational persistence generated by parental networks and preferences transmission may be detrimental to welfare. Furthermore, we show that search frictions interact with parental networks, amplifying their importance and their adverse effects on the aggregate equilibrium.

We evaluate the cost of increasing unemployment benefits and find that the restricted model understates the cost by a factor of two. Hence, we conclude that modeling multiple sources of intergenerational transmission is crucial not only to understanding the consequences of persistence, but also for the assessment of labor market policy.

Interesting directions for future research are the study of the cross–gender patterns of occupational persistence, and asymmetric equilibria across occupations. Analyzing the latter would make it possible to capture heterogeneity across occupations, though it requires much richer data in order to reliably estimate the separate channels of persistence at the occupational level.

References

Abbott, Brant, Gallipoli, Giovanni, Meghir, Costas, & Violante, Giovanni L. 2013. *Education policy and intergenerational transfers in equilibrium.* Tech. rept. National Bureau of Economic Research.

Acemoglu, Daron, & Shimer, Robert. 2000. Productivity gains from unemployment insurance. *European Economic Review*, **44**(7), 1195–1224.

Aina, Carmen, & Nicoletti, Cheti. 2014. The intergenerational transmission of liberal professions: nepotism versus abilities. *University of York Discussion Papers in Economics.*

Andres, Lesley, Anisef, Paul, Krahn, Harvey, Looker, Dianne, & Thiessen, Victor. 1999. The persistence of social structure: Cohort, class and gender effects on the occupational aspirations and expectations of Canadian youth. *Journal of youth studies*, **2**(3), 261–282.

Arulampalam, Wiji. 2001. Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal*, **111**.

Barczyk, Daniel, & Kredler, Matthias. 2014. *Evaluating long-term care policy options, taking the family seriously.* Tech. rept. Working paper.

Becker, Gary S, & Tomes, Nigel. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *The Journal of Political Economy*, 1153–1189.

Bentolila, Samuel, Michelacci, Claudio, & Suarez, Javier. 2010. Social Contacts and Occupational Choice. *Economica.*

Bisin, Alberto, & Verdier, Thierry. 2005. Cultural transmission. *The New Palgrave Dictionary of Economics.*

Björklund, Anders, & Jäntti, Markus. 2009. Intergenerational income mobility and the role of family background. *Oxford Handbook of Economic Inequality, Oxford University Press, Oxford.*

Black, Sandra E, & Devereux, Paul J. 2010. *Recent developments in intergenerational mobility.* Tech. rept. National Bureau of Economic Research.

Blau, Peter M, & Duncan, Otis Dudley. 1967. *The American occupational structure.* ERIC.

Calvó-Armengol, Antoni, & Jackson, Matthew O. 2007. Networks in labor markets: Wage and employment dynamics and inequality. *Journal of economic theory*, **132**(1), 27–46.

Caner, Asena, & Okten, Cagla. 2010. Risk and career choice: Evidence from Turkey. *Economics of Education Review*, **29**(6), 1060–1075.

Carrillo-Tudela, Carlos, & Visschers, Ludo. 2014. Unemployment and Endogenous Reallocation over the Business Cycle. *Working Paper*.

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

Celik, Murat Alp. 2015. *Does the Cream Always Rise to the Top? The Misallocation of Talent and Innovation*. Tech. rept. Mimeo.

Checchi, Daniele. 1997. Education and intergenerational mobility in occupations: a comparative study. *American journal of Economics and Sociology*, 331–351.

Chevalier, Arnaud, Denny, Kevin, & McMahon, Dorren. 2009. Intergenerational mobility and education equality. *Education and inequality across Europe*, 260–281.

Cingano, Federico, & Rosolia, Alfonso. 2012. People I know: job search and social networks. *Journal of Labor Economics*, **30**(2), 291–332.

Constant, Amelie F, & Zimmermann, Klaus F. 2004. Occupational choice across generations. *CEPR Discussion Paper*.

Corak, Miles. 2006. Do poor children become poor adults? Lessons from a cross country comparison of generational earnings mobility. *Research on Economic Inequality*, **13**.

Corak, Miles, & Piraino, Patrizio. 2011. The intergenerational transmission of employers. *Journal of Labor Economics*, **29**(1), 37–68.

Di Pietro, Giorgio, & Urwin, Peter. 2003. Intergenerational mobility and occupational status in Italy. *Applied Economics Letters*, **10**(12), 793–797.

Doepke, Matthias, & Zilibotti, Fabrizio. 2008. Occupational choice and the spirit of capitalism. *The Quarterly Journal of Economics*, **123**(2), 747–793.

Dohmen, Thomas, Falk, Armin, Huffman, David, & Sunde, Uwe. 2011. The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, **79**(2), 645–677.

Dustmann, Christian. 2004. Parental background, secondary school track choice, and wages. *Oxford Economic Papers*.

Dustmann, Christian, Glitz, Albrecht, & Schönberg, Uta. forthcoming. Referral-based job search networks. *Review of Economic Studies*.

Eberharter, Veronika B. 2008. Parental Background and Intergenerational Occupational Mobility - Germany and the United States compared. *The Journal of Income Distribution*, **17**(2).

Ermisch, John, & Francesconi, Marco. 2002. Intergenerational mobility in Britain: new evidence from the BHPS. *Generational income mobility in North America and Europe (Cambridge University Press: Cambridge)*.

Ermisch, John, Jantti, Markus, & Smeeding, Timothy M. 2012. *From parents to children: The intergenerational transmission of advantage*. Russell Sage Foundation.

Escriche, Luisa. 2007. Persistence of Occupational Segregation: the Role of the Intergenerational Transmission of Preferences*. *The Economic Journal*, **117**(520), 837–857.

Galenianos, Manolis. 2014. Hiring through referrals. *Journal of Economic Theory*, **152**, 304–323.

Gayle, George-Levi, Golan, Limor, & Soytas, Mehmet A. 2015. What is the source of the intergenerational correlation in earnings? *Working Paper*.

Golosov, Mikhail, Maziero, Pricila, & Menzio, Guido. 2013. Taxation and redistribution of residual income inequality. *Journal of Political Economy*.

Granovetter, Mark S. 1973. The strength of weak ties. *American journal of sociology*, 1360–1380.

Groes, Fane, Kircher, Philipp, & Manovskii, Iourii. 2014. The U-shapes of occupational mobility. *The Review of Economic Studies*, **82**(2), 659–692.

Güell, Maia, Pellizzari, Michele, Pica, Giovanni, & Mora, José Vicente Rodriguez. 2015. Correlating Social Mobility and Economics Outcomes. *Working Paper*.

Hellerstein, Judith K, & Morrill, Melinda Sandler. 2011. Dads and daughters the changing impact of fathers on women's occupational choices. *Journal of Human Resources*, **46**(2), 333–372.

HENSVIK, LENA, & SKANS, OSKAR NORDSTROM. 2013. Social networks, employee selection and labor market outcomes: Toward an empirical analysis. *Journal of Labor Economics, forthcoming.*

HERTZ, TOM. 2006. Understanding mobility in America. *Center for American Progress Discussion Paper.*

HOBIJN, BART, & ŞAHIN, AYŞEGÜL. 2009. Job-finding and separation rates in the OECD. *Economics Letters*, **104**(3), 107–111.

HORVÁTH, GERGELY. 2014. Occupational mismatch and social networks. *Journal of Economic Behavior & Organization*, **106**, 442–468.

HOUT, MICHAEL, & BELLER, EMILY. 2006. Intergenerational social mobility: The United States in comparative perspective. *The Future of Children*, **16**(2), 19–36.

HSIEH, CHANG-TAI, HURST, ERIK, JONES, CHARLES I, & KLENOW, PETER J. 2013. *The allocation of talent and us economic growth.* Tech. rept. National Bureau of Economic Research.

KAMBOUROV, GUERGOUI, & MANOVSKII, IOURII. 2009. Occupational Mobility and Wage Inequality. *The Review of Economic Studies*, **76**(2).

KEANE, MICHAEL P., & WOLPIN, KENNETH I. 1997. The Career Decisions of Young Men. *Journal of Political Economy*, **105**(3).

KRAMARZ, FRANCIS, & SKANS, OSKAR NORDSTRÖM. 2014. When strong ties are strong: Networks and youth labour market entry. *The Review of Economic Studies*, **81**(3), 1164–1200.

LABAND, DAVID N, & LENTZ, BERNARD F. 1983. Like father, like son: Toward an economic theory of occupational following. *Southern Economic Journal*, 474–493.

LEE, SANG YOON, & SESHADRI, ANANTH. 2014. On the intergenerational transmission of economic status. *Unpublished manuscript, University of Wisconsin–Madison, Department of Economics.*

LENTZ, BERNARD F, & LABAND, DAVID N. 1989. Why so many children of doctors become doctors: Nepotism vs. human capital transfers. *Journal of Human Resources*, 396–413.

LO BELLO, SALVATORE, & MORCHIO, IACOPO. 2017. Parental Links and Labor Market Outcomes: Evidence from the UK. *Working paper.*

Long, Jason, & Ferrie, Joseph. 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. *The American Economic Review*, **103**(4), 1109–1137.

McCall, Brian P. 1991. A Dynamic Model of Occupational Choice. *Journal of Economic Dynamics and Control*.

Miller, Robert A. 1984. Job Matching and Occupational Choice. *Journal of Political Economy*, **92**(6).

Montgomery, James D. 1991. Social networks and labor-market outcomes: Toward an economic analysis. *The American economic review*, **81**(5), 1408–1418.

Mora, José V Rodríguez. 2009. *The misallocation of talent*. CREI, Centre de Recerca en Economia Internacional.

Munshi, Kaivan, & Rosenzweig, Mark. 2016. Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *The American Economic Review*, **106**(1), 46–98.

Paola, Maria. 2013. The determinants of risk aversion: the role of intergenerational transmission. *German Economic Review*, **14**(2), 214–234.

Papageorgiou, Theodore. 2014. Learning your comparative advantages. *The Review of Economic Studies*, **81**(3), 1263–1295.

Pellizzari, Michele. 2010. Do friends and relatives really help in getting a good job? *Industrial & Labor Relations Review*, **63**(3), 494–510.

Pellizzari, Michele, Pica, Giovanni, *et al.* 2011. Liberalizing professional services: Evidence from italian lawyers. *U. of Bocconi working paper (May 2011)*.

Petrongolo, Barbara, & Pissarides, Christopher A. 2001. Looking into the black box: A survey of the matching function. *Journal of Economic literature*, **39**(2), 390–431.

Restuccia, Diego, & Urrutia, Carlos. 2004. Intergenerational Persistence of Earnings; The Role of Early and College Education. *The American Economic Review*, **94**(5).

Sinha, Rishabh. 2014. Intergenerational Occupational Mobility and Labor Productivity. *Working Paper*.

Solon, Gary. 1992. Intergenerational income mobility in the United States. *The American Economic Review*, 393–408.

Solon, Gary. 2002. Cross-country differences in intergenerational earnings mobility. *The Journal of Economic Perspectives*, **16**(3), 59–66.

Stier, Haya, & Grusky, David B. 1990. An overlapping persistence model of career mobility. *American Sociological Review*, 736–756.

Topa, Giorgio. 2001. Social interactions, local spillovers and unemployment. *The Review of Economic Studies*, **68**(2), 261–295.

Wiczer, David. 2014. Long Term Unemployment: Attached and Mismatched? *Working Paper*.

Appendix A Derivations and Proofs

Empirical Prediction 2

Using the symmetry of the equilibrium, we can write the following:

$$P(\bar{w}=1|\bar{w}^F=1)=P(\bar{w}=1|\bar{w}^F=1;o^F=A). \tag{56}$$

Next, we define $\bar{w}_j$ as a dummy taking value 1 when the joint event $(\bar{w}=1;o=j)$ is satisfied (that is, the individual is a high–earner in occupation j). We can now rewrite the previous expression as:

$$P(\bar{w}=1|\bar{w}^F_A=1)=P(\bar{w}_A=1|\bar{w}^F_A=1)+P(\bar{w}_B=1|\bar{w}^F_A=1), \tag{57}$$

using the fact that $\bar{w}_A=1$ and $\bar{w}_B=1$ are mutually exclusive events. The first term of Equation (57) can be calculated as follows:

$$\begin{aligned} P(\bar{w}_A=1|\bar{w}^F_A=1) &= P(o=A;\tau=A|o^F=A;\tau^F=A)= \\ &= P(o=A|\tau=A;o^F=A;\tau^F=A)P(\tau=A|o^F=A;\tau^F=A)=\rho. \end{aligned} \tag{58}$$

In contrast, the second term of Equation (57) can be rewritten as follows:

$$\begin{aligned} P(\bar{w}_B=1|\bar{w}^F_A=1) &= P(o=B;\tau=B|o^F=A;\tau^F=A)= \\ &= P(o=B|\tau=B;o^F=A;\tau^F=A)P(\tau=B|o^F=A;\tau^F=A) \\ &= (1-\mu)(1-\rho). \end{aligned} \tag{59}$$

By substituting the last two equations into (57), we arrive at the result.

Lemma 1

Proof. Taking the derivative of m^* w.r.t. μ:

$$\frac{\partial m^*}{\partial \mu}=\frac{-\rho(1+\mu(1-2\rho))-(1-\mu\rho)(1-2\rho)}{[1+\mu(1-2\rho)]^2}$$

The numerator can be simplified to $(1+\mu)(\rho-1)$, which is clearly non–positive, $\forall\ \rho\leq 1$. Therefore, $\frac{\partial m^*}{\partial \mu}\leq 0$.

Taking the derivative of m^* w.r.t. ρ:

$$\frac{\partial m^*}{\partial \rho} = \frac{-\mu(1+\mu(1-2\rho)) - (1-\mu\rho)(-2\mu)}{[1+\mu(1-2\rho)]^2}$$

The numerator can be simplified to $(1+\mu)\mu$, which is clearly non–negative, $\forall\ \mu \geq 0$. Therefore, $\frac{\partial m^*}{\partial \rho} \geq 0$. □

Lemma 2

Proof. Taking the derivative of Ψ^* w.r.t. μ:

$$\frac{\partial \Psi^*}{\partial \mu} = \frac{(1-2\rho)[1+\mu(1-2\rho) - \rho - \mu + 2\mu\rho]}{[1+\mu(1-2\rho)]^2}$$

The numerator can be simplified to $(1-2\rho)(1-\rho)$, which is clearly non–positive, $\forall\ \rho \geq \frac{1}{2}$. Therefore, $\frac{\partial \Psi^*}{\partial \mu} \leq 0$.
Taking the derivative of Ψ^* w.r.t. ρ:

$$\frac{\partial \Psi^*}{\partial \rho} = \frac{(1-2\mu)[1+\mu(1-2\rho)] + 2\mu[\rho+\mu-2\rho\mu]}{[1+\mu(1-2\rho)]^2}$$

The numerator can be simplified to $1-\mu$, which is clearly non–negative, $\forall\ \mu \leq 1$. Therefore, $\frac{\partial \Psi^*}{\partial \rho} \geq 0$. □

Lemma 3

Proof. Taking the derivative of γ^* w.r.t. μ:

$$\frac{\partial \gamma^*}{\partial \mu} = -\frac{[(1-\Psi^*)(1+N) - \mu(1+N)\frac{\partial \Psi^*}{\partial \mu}U^* - (1-\Psi^*)\mu(1+N)\frac{\partial U^*}{\partial \mu}}{(U^*)^2}$$

The numerator can be rearranged as
$-\left[(1-\Psi)[(1+N)U^* - \mu(1+N)\frac{\partial U^*}{\partial \mu}] - \mu(1+N)\frac{\partial \Psi^*}{\partial \mu}U^*\right]$.
The term $\mu(1+N)\frac{\partial \Psi^*}{\partial \mu}U^*$ is clearly non–positive, since $\frac{\partial \Psi^*}{\partial \mu} \leq 0$. For the first term, it suffices to show that: $U^* \geq \mu\frac{\partial U^*}{\partial \mu}$. The left–hand side of this inequality can be written as $1+\mu N + \Psi(1-\mu)n$, while the right–hand side is: $\mu N + \mu\frac{\partial \Psi^*}{\partial \mu}(1-\mu)n - \mu\Psi n$. Cancelling terms yields: $1+\Psi^* n \geq \mu\frac{\partial \Psi^*}{\partial \mu}(1-\mu)n$, which is always satisfied, since $\frac{\partial \Psi^*}{\partial \mu} \leq 0$. Hence, the whole term in square brackets is non-negative, and therefore $\frac{\partial \gamma^*}{\partial \mu} \leq 0$.

Taking the derivative of γ^* w.r.t. ρ yields:

$$\frac{\partial \gamma^*}{\partial \rho} = -\frac{-\mu(1+N)\frac{\partial \Psi^*}{\partial \rho} - (1-\Psi^*)\mu(1+N)\frac{\partial \Psi^*}{\partial \rho}[(1-\mu)n]}{(U^*)^2}$$

The numerator can be rearranged to: $\frac{\partial \Psi^*}{\partial \rho}[-\mu(1+N) - (1-\Psi^*)\mu(1+N)(1-\mu)n]$. The term in square brackets is non–positive, and the fact that $\frac{\partial \Psi^*}{\partial \rho} \geq 0$ completes the claim that $\frac{\partial \gamma^*}{\partial \rho} \geq 0$. □

Lemma 4

We evaluate the free–entry condition at equilibrium:

$$q(\theta^*)(1-\beta)(1+\gamma^* a) = \kappa$$

The fact that $q'(\theta) < 0$ implies that θ^* adjusts to the new equilibrium in the same direction as γ^*. The fact that $\frac{\partial \gamma^*}{\partial \mu} \leq 0$, $\frac{\partial \gamma^*}{\partial \rho} \geq 0$ (shown in Lemma 3) completes the claim that $\frac{\partial \theta^*}{\partial \mu} \leq 0$ and $\frac{\partial \theta^*}{\partial \rho} \geq 0$.

Proposition 1

We evaluate persistence at equilibrium:

$$\mathcal{P}^* = \mu + (1-\mu)\Psi^*$$

Taking the derivative of $\mathcal{P}^*$ w.r.t. μ:

$$\frac{\partial \mathcal{P}^*}{\partial \mu} = (1-\mu)\frac{\partial \Psi^*}{\partial \mu} + (1-\Psi^*)$$

The first term is negative, while the second is always positive. We can show that even in the case of $\mu = 0$ (that is, making the negative term as large as possible), the sum of the two is still positive. To see this, we use some of the expressions derived earlier:

$$(1-\Psi^*) = \frac{1-\rho}{[1+\mu(1-2\rho)]} \geq -\frac{(1-2\rho)(1-\rho)}{[1+\mu(1-2\rho)]^2} = -\frac{\partial \Psi^*}{\partial \mu},$$

where the inequality follows from the fact that $1 \geq -(1-2\rho)(1-\mu)$. Hence, $\frac{\partial \mathcal{P}^*}{\partial \mu} \geq 0$. We now take the derivative of $\mathcal{P}^*$ w.r.t. ρ:

$$\frac{\partial \mathcal{P}^*}{\partial \rho} = (1-\mu)\frac{\partial \Psi^*}{\partial \rho},$$

which is clearly positive, given that $\frac{\partial \Psi^*}{\partial \rho} \geq 0$.

Definition of different subgroups of workers

Subgroup #	Father Well–sorted	Same τ as father	Large Network	Share
1	Yes	Yes	Yes	$m\rho\mu$
2	Yes	Yes	No	$m\rho(1-\mu)$
3	Yes	No	Yes	$m(1-\rho)\mu$
4	Yes	No	No	$m(1-\rho)(1-\mu)$
5	No	Yes	Yes	$(1-m)\rho\mu$
6	No	Yes	No	$(1-m)\rho(1-\mu)$
7	No	No	Yes	$(1-m)(1-\rho)\mu$
8	No	No	No	$(1-m)(1-\rho)(1-\mu)$

Table XI. Description of the different subgroups of workers.

Expression for m and $\mathcal{P}$

$$m(m_1, ..., m_8) = \frac{\rho\mu m_5 + \rho(1-\mu)m_6 + (1-\rho)\mu m_7 + (1-\rho)(1-\mu)m_8}{1 + \rho\mu(m_5 - m_1) + \rho(1-\mu)(m_6 - m_2) + (1-\rho)\mu(m_7 - m_3) + (1-\rho)(1-\mu)(m_8 - m_4)}.$$

$$\begin{aligned} \mathcal{P}(m, m_1, ..., m_8) = m\Big[&\rho(\mu m_1 + (1-\mu)m_2) + (1-\rho)(\mu(1-m_3) + (1-\mu)(1-m_4))\Big] \quad (60) \\ &+ (1-m)[\rho(\mu(1-m_5) + (1-\mu)(1-m_6)) + (1-\rho)(\mu m_7 + (1-\mu)m_8)\Big]. \end{aligned}$$

$$\frac{\partial \mathcal{W}^{SP}}{\partial m}\frac{\partial m}{\partial m_j}\Big|_{\vec{m}^{SE}} = \Big[[p(\theta)y(1+a) - \kappa\theta](2\rho-1)(1-\mu)n + p(\theta)ya(2\rho-1)\mu(1+N)\Big]\frac{\partial m}{\partial m_j}\Big|_{\vec{m}^{SE}} > 0. \quad (61)$$

Proposition 2: *Depending on the parameter values, the SE allocation $\vec{m}^{SE}$ does not necessarily coincide with the efficient one.*

- *Under $a > \frac{N}{1+N} - \frac{N}{1+N}\frac{\kappa\theta}{p(\theta)y}$ (Condition a/N^*_+) and $a > \frac{n}{1-n} - \frac{n}{1-n}\frac{\kappa\theta}{p(\theta)y}$ (Condition a/n^{**}_+): $m_3^{*,SP}, m_5^{*,SP} > 0$; $m_4^{*,SP} = m_6^{*,SP} = 1$.*
  That is, the SP wants to achieve less mismatch than in the SE.

- *Under* $a < \frac{N}{1+N} - \frac{N}{1+N}\frac{\kappa\theta}{p(\theta)y}$ *(Condition* a/N_{-}^{*}*) and* $a < \frac{n}{1-n} - \frac{n}{1-n}\frac{\kappa\theta}{p(\theta)y}$ *(Condition* a/n_{-}^{**}*),* $\exists\bar{\rho} > \frac{1}{2}$ *such that* $\forall\rho < \bar{\rho}$ $: m_3^{*,SP} = m_5^{*,SP} = 0$*;* $m_4^{*,SP}, m_6^{*,SP} < 1$*.*
  That is, the SP wants to achieve more mismatch than in the SE.

Proof. First notice that, under $\vec{m}^{SE} = \{1,1,0,1,0,1,1,1\}$, the *reallocation effect* is always strictly positive:

$$\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m}\frac{\partial m}{\partial m_j}\Big|_{\vec{m}^{SE}} = \Big[[p(\theta)y(1+a)-\kappa\theta](2\rho-1)(1-\mu)n+p(\theta)ya(2\rho-1)\mu(1+N)\Big]\frac{\partial m}{\partial m_j}\Big|_{\vec{m}^{SE}} > 0. \tag{62}$$

Intuitively, the strength of this effect depends on the size of ρ. The larger ρ is, the more scope the SP has to sort workers according to productivity, so that fewer workers have a tradeoff in the next period. It can be seen that as $\rho \to \frac{1}{2}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m}\frac{\partial m}{\partial m_j}\Big|_{\vec{m}^{SE}} \to 0$. When productive types are nearly independent across generations, the SP can align productive advantage and networks to a very limited extent (in the limit, he cannot do so at all).

Inspection of the *inner effect*, evaluated at the SE allocation, reveals that $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_1}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_2}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_7}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_8} > 0$. This implies that $m_1^{*,SP} = m_2^{*,SP} = m_7^{*,SP} = m_8^{*,SP} = 1$. It remains to find the optimal values of m_3,m_4, m_5 and m_6.

It is important to remember that under the SE allocation, the *reallocation effect (RE)* is positive.
The first statement of Proposition 2 follows from the fact that $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_3}\Big|_{\vec{m}^{SE}}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_5}\Big|_{\vec{m}^{SE}} > 0$ under Condition (a/N_{+}^{*}), along with $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_4}\Big|_{\vec{m}^{SE}}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_6}\Big|_{\vec{m}^{SE}} > 0$ under Condition (a/n_{+}^{**}).
The second statement of Proposition 2 follows from the fact that $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_3}\Big|_{\vec{m}^{SE}}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_5}\Big|_{\vec{m}^{SE}} < 0$ under Condition (a/N_{-}^{*}), along with $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_4}\Big|_{\vec{m}^{SE}}$, $\frac{\partial \mathcal{W}^{\mathcal{SP}}}{\partial m_6}\Big|_{\vec{m}^{SE}} < 0$ under Condition (a/n_{-}^{**}) and the fact that the $RE \to 0$ as $\rho \to \frac{1}{2}$.
Note that these are only sufficient conditions. Given that the RE is positive, there will be additional regions of the parameter space in which the statements are true. □

Proposition 3

We focus on equilibria in which $\eta_V(\theta) \geq (1-\beta)$, so that the indirect effect (GE effect) of an increase in μ (ρ) is negative (positive). We provide a sufficient condition for the direct effect of an increase in μ (ρ) to also be negative (positive). The direct effect of μ is given by:

$$\frac{\partial \mathcal{W}}{\partial \mu} = 2\left[p(\theta)(1+\gamma a)y - \kappa\theta\right]\frac{\partial U}{\partial \mu} + 2p(\theta)ayU\frac{\partial \gamma}{\partial \mu}$$

Replacing $\frac{\partial \gamma}{\partial \mu}$ with the expression found in Lemma 3 and regrouping terms yields:

$$\left[U^2(1+\gamma a)\beta + a(1-\Psi)\mu(1+N)\right]\frac{\partial U}{\partial \mu} \leq a\left[(1-\Psi)(1+N) - \mu(1+N)\frac{\partial \Psi}{\partial \mu}\right]$$

Replacing $\frac{\partial U}{\partial \mu} = N + \frac{\partial \Psi}{\partial \mu}(1-\mu)n - \Psi n$ and defining $A \equiv [U^2(1+\gamma a)\beta + a(1-\Psi)\mu(1+N)]$, we get:

$$-a(1-\Psi)(1+N) \leq -a\mu(1+N)\frac{\partial \Psi}{\partial \mu} - A\left[(1-\mu)n\frac{\partial \Psi}{\partial \mu} + (\Psi n - N)\right]$$

It can easily be verified that $(1-\mu)n\frac{\partial \Psi}{\partial \mu} \leq N - \Psi n$ if the following condition is satisfied:

$$\frac{N}{n} \leq \frac{(1-2\rho)(1-\rho)(1-\mu)n}{[1+\mu(1-2\rho)]^2} + \frac{\rho+\mu-2\rho\mu}{[1+\mu(1-2\rho)]}$$

Hence, $\frac{d\mathcal{W}}{d\mu} \leq 0$.
In the text we have already shown that $\frac{\partial \mathcal{W}}{\partial \rho}$ has to be non–negative. As a consequence, $\frac{d\mathcal{W}}{d\rho} \geq 0$, since it is the sum of two positive components.

Appendix B Worker Flows

The evolution of the stock of unemployed and employed workers is the result of optimal relocation decisions, age shocks and labor market shocks (creation of new matches and destruction of existing ones). Define $g^k_{\Omega}(\tilde{\Omega}) = P(\Omega' = \Omega|\tilde{\Omega})$ to be the probability measure that a worker of type $\tilde{\Omega}$ with employment status k changes to type Ω in the following period. This probability is defined over the multidimensional distribution of Ω. In particular, it involves changes in: the temporary preference vectors (his own or his father's), occupation–specific human capital and networks stocks (his own or his father's), father's occupation or employment status.

Define $u'_{o,F}(\Omega)$ to be the subsequent period's measure of unemployed fathers of type Ω in occupation o:

$$\begin{aligned} u'_{o,F}(\Omega) = & \int_{\Omega} \Big[\hat{u}_{o,F}(\tilde{\Omega}) \, (1 - R^U_{o,F}(\tilde{\Omega})) \, (1 - p_{o,F}(\tilde{\Omega})) \, g^U_{\Omega}(\tilde{\Omega}) \\ & + \hat{e}_{o,F}(\tilde{\Omega}) \, (1 - R^E_{o,F}(\tilde{\Omega})) \, \delta \, g^E_{\Omega}(\tilde{\Omega}) \Big] \, d\tilde{\Omega} \\ & + \sum_{\tilde{o} \neq o} \int_{\Omega} \Big[\hat{u}_{\tilde{o},F}(\tilde{\Omega}) \, R^U_{\tilde{o},F}(\tilde{\Omega}) \, + \hat{e}_{\tilde{o},F}(\tilde{\Omega}) \, R^E_{\tilde{o},F}(\tilde{\Omega}) \Big] \\ & \cdot \mathbb{1}\{ j^*_F(\tilde{\Omega}) = o \} \, (1 - p_{o,F}(\tilde{\Omega})) \, g^U_{\Omega}(\tilde{\Omega}) \, d\tilde{\Omega}, \end{aligned} \tag{63}$$

where $\hat{u}_{o,F} = u_{o,F} \, (1 - \zeta) + u_{o,S} \, \zeta$, and $\hat{e}_{o,F} = e_{o,F} \, (1 - \zeta) + e_{o,S} \, \zeta$. These are the measures of workers after the age shock, that is they include fathers who did not die, as well as sons who became fathers.

Equation (63) is composed of four different terms: the first two refer respectively to unemployed workers in occupation o who decided not to relocate and did not find a job, and employed workers in occupation o who did not relocate and lost their job; the last two are (unemployed and employed) workers who decided to relocate into occupation o but did not find a job in the previous period.

For employed fathers, $e'_o(\Omega)$ is defined as:

$$
\begin{aligned}
e'_{o,F}(\Omega) = \int_{\Omega} \Big[& \hat{e}_{o,F}(\tilde{\Omega})\,(1 - R^{E}_{o,F}(\tilde{\Omega}))\,(1-\delta)\,g^{E}_{\Omega}(\tilde{\Omega}) \\
& + \hat{u}_{o,F}(\tilde{\Omega})\,(1 - R^{U}_{o,F}(\tilde{\Omega}))\,p_{o,F}(\tilde{\Omega})\,g^{U}_{\Omega}(\tilde{\Omega}) \Big]\,d\tilde{\Omega} \\
& + \sum_{\tilde{o}\neq o} \int_{\Omega} \Big[\hat{u}_{\tilde{o},F}(\tilde{\Omega})\,R^{U}_{\tilde{o},F}(\tilde{\Omega}) + \hat{e}_{\tilde{o},F}(\tilde{\Omega})\,R^{E}_{\tilde{o},F}(\tilde{\Omega}) \Big] \\
& \cdot \mathbb{1}\{j^{*}_{F}(\tilde{\Omega}) = o\}\,p_{o,F}(\tilde{\Omega})\,g^{U}_{\Omega}(\tilde{\Omega})\,d\tilde{\Omega}.
\end{aligned}
\tag{64}
$$

The stock of employed is made up of workers who were already employed in the previous period in the same occupation and did not lose their job nor did they find it profitable to relocate, and the mass of unemployed workers who did not want to relocate and found a vacancy, plus all workers who had just relocated into occupation o and found a job.

The distribution of employed sons, is exactly symmetric to that of the fathers:

$$
\begin{aligned}
e'_{o,S}(\Omega) = \int_{\Omega} \Big[& (1-\zeta)e_{o,S}(\tilde{\Omega})\,(1 - R^{E}_{o,S}(\tilde{\Omega}))\,(1-\delta)\,g^{E}_{\Omega}(\tilde{\Omega}) \\
& + (1-\zeta)u_{o,S}(\tilde{\Omega})\,(1 - R^{U}_{o,S}(\tilde{\Omega}))\,p_{o,S}(\tilde{\Omega})\,g^{U}_{\Omega}(\tilde{\Omega}) \Big]\,d\tilde{\Omega} \\
& + \sum_{\tilde{o}\neq o} \int_{\Omega} \Big[(1-\zeta)u_{\tilde{o},S}(\tilde{\Omega})\,R^{U}_{\tilde{o},S}(\tilde{\Omega}) + (1-\zeta)e_{\tilde{o},S}(\tilde{\Omega})\,R^{E}_{\tilde{o},S}(\tilde{\Omega}) \Big] \\
& \cdot \mathbb{1}\{j^{*}_{S}(\tilde{\Omega}) = o\}\,p_{o,S}(\tilde{\Omega})\,g^{U}_{\Omega}(\tilde{\Omega})\,d\tilde{\Omega}.
\end{aligned}
\tag{65}
$$

Finally, the distribution of unemployed sons is as follows:

$$
\begin{aligned}
u'_{o,S}(\Omega) = \int_{\Omega} \Big[& (1-\zeta)u_{o,S}(\tilde{\Omega})\,(1 - R^{U}_{o,S}(\tilde{\Omega}))\,(1 - p_{o,S}(\tilde{\Omega}))\,g^{U}_{\Omega}(\tilde{\Omega}) \\
& + (1-\zeta)e_{o,S}(\tilde{\Omega})\,(1 - R^{E}_{o,S}(\tilde{\Omega}))\,\delta\,g^{E}_{\Omega}(\tilde{\Omega}) \Big]\,d\tilde{\Omega} \\
& + \sum_{\tilde{o}\neq o} \int_{\Omega} \Big[(1-\zeta)u_{\tilde{o},S}(\tilde{\Omega})\,R^{U}_{\tilde{o},S}(\tilde{\Omega}) + (1-\zeta)e_{\tilde{o},S}(\tilde{\Omega})\,R^{E}_{\tilde{o},S}(\tilde{\Omega}) \Big] \\
& \cdot \mathbb{1}\{j^{*}_{S}(\tilde{\Omega}) = o\}\,(1 - p_{o,S}(\tilde{\Omega}))\,g^{U}_{\Omega}(\tilde{\Omega})\,d\tilde{\Omega} \\
& + \zeta\,\frac{\mathbb{1}\{\Omega \in \Omega^{NB}\}}{\int_{\Omega} \mathbb{1}\{\Omega \in \Omega^{NB}\}},
\end{aligned}
\tag{66}
$$

with the only difference being the last term, which represents the flow of newborns, randomly directed to the subset Ω^{NB} of the entire state space.

Appendix C Other Figures and Tables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Son is in:	Occ. 1	Occ. 2	Occ. 3	Occ. 4	Occ. 5	Occ. 6	Occ. 7	Occ. 8	Occ. 9
Father is in:									
Occ. 1	0.0335^{***}								
	(0.003)								
Occ. 2		0.0365^{***}							
		(0.003)							
Occ. 3			0.0550^{***}						
			(0.005)						
Occ. 4				0.0411^{***}					
				(0.005)					
Occ. 5					0.151^{***}				
					(0.004)				
Occ. 6						0.0398^{***}			
						(0.005)			
Occ. 7							0.0159^{***}		
							(0.006)		
Occ. 8								0.127^{***}	
								(0.003)	
Occ. 9									0.105^{***}
									(0.006)
N	62114	62114	62114	62114	62114	62114	62114	62114	62114
R^2	0.073	0.196	0.079	0.047	0.106	0.053	0.076	0.084	0.089

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

Table XII. Regressions of Occupational Choice (dummy that takes a value of 1 if the offspring is in a given occupation, 0 otherwise); coefficient for **father in a given occupation (dummy variable, 0 if the father is in some other occupation)**, standard errors. All models are linear probability models, and include a third-degree polynomial in age and dummies for education, region of residence, smoking behavior, marital status, ethnicity and quarter. Occupational codes are as defined in Table I: 1) Managers & Administrators; 2) Professional; 3) Associate Professional; 4) Clerical & Secretarial; 5) Craft & Related; 6) Personal & Protective Service; 7) Sales; 8) Plant & Machine; 9) Agriculture & Elementary. Source: BHPS (1991–2008).

Occ. code	Occupational group	Likelihood Ratio	Occ. Share
11	Production Managers in Manuf., Construction	6.23	0.0133
12	Specialist Managers	2.03	0.0210
13	Office Managers	0.00	0.0129
14	Managers in Transport and Storing	4.69	0.0084
16	Managers in Farming	33.05	0.0084
17	Managers in Service Industry	1.92	0.0301
19	Managers and Administrators NEC	3.04	0.0083
21	Engineers and Technologists	4.79	0.0144
22	Health Professionals	0.00	0.0064
23	Teaching Professionals	1.67	0.0112
25	Business and Financial Professionals	12.19	0.0102
31	Draughtspersons	1.77	0.0109
32	Computer Analyst/Programmers	0.00	0.0301
34	Health Associate Professionals	10.66	0.0066
36	Business and Financial Associate Professionals	6.23	0.0186
37	Social Welfare Associate Professionals	0.00	0.0051
38	Literary, Artistic and Sports Professionals	3.01	0.0264
39	Associate Professionals and Technical Occ.s NEC	2.52	0.0056
40	Administrative/Clerical Officers	2.71	0.0088
41	Numerical Clerks and Cashiers	0.00	0.0380
42	Filing and Record Clerks	2.05	0.0188
43	Clerks	0.82	0.0285
44	Stores and Despatch Clerks	3.16	0.0276
50	Construction Trades	5.81	0.0424
51	Metal Machining	1.64	0.0363
52	Electrical/Electronic Trades	6.43	0.0541
53	Metal Forming, Welding and Related	2.57	0.0360
54	Vehicle Traders	5.78	0.0317
57	Woodworking Trades	7.81	0.0322
58	Food Preparation Trades	29.71	0.0103
59	Other Craft and Related Occupations NEC	0.50	0.0185
61	Security and Protective Service	5.08	0.0059
62	Catering Occupations	3.20	0.0268
71	Sales Representatives	1.72	0.0166
72	Sales Assistants and Check-out Operators	0.28	0.0534
80	Food, Drink and Tobacco Process Operatives	35.81	0.0086
82	Chemicals, Paper, Plastics Operatives	5.42	0.0120
84	Metal Working Process Operatives	3.82	0.0088
85	Assemblers/Lineworkers	5.16	0.0132
86	Other Routine Process Operatives	3.51	0.0148
87	Road Transport Operatives	4.54	0.0320
88	Other Transport and Machinery Operatives	7.25	0.0055
89	Plant and Machine Operatives NEC	4.09	0.0138
90	Other Occ.s in Agriculture, Forestry and Fishing	16.96	0.0108
92	Other Occ.s in Construction	11.78	0.0096
94	Other Occ.s in Communication	0.42	0.0080
95	Other Occ.s in Sales and Services	9.22	0.0336
99	Other Occ.s NEC	0.45	0.0120
	Average (unweighted)	5.69	
	Average (weighted)	4.71	

Table XIII. Occupational Persistence (Likelihood Ratios), 2–digit level. The table presents the likelihood ratios for occupations in which at least 0.5% of the workforce are employed, due to the limited size of the sample. Averages are taken with respect to all occupations, including the ones not reported in the table. Source: BHPS (1991–2008).

Occupational sector	Likelihood Ratio		
	Bottom 33%	Mid 33%	Top 33%
Managers & Administrators	1.49	0.82	1.51
Professional	2.93	2.61	2.35
Associate Professional & Technical	1.40	1.84	1.60
Clerical & Secretarial	1.43	1.45	1.01
Craft & Related	1.64	1.57	1.46
Personal & Protective Service	1.81	2.72	0.50
Sales	0.99	1.56	1.47
Plant & Machine	2.49	1.48	1.88
Agriculture & Elementary	3.02	2.31	2.72
Average (unweighted)	1.91	1.82	1.61
Average (weighted)	1.79	1.63	1.65

Table XIV. Occupational Persistence Indexes, by Father's Income (Likelihood Ratios). Source: BHPS (1991–2008).

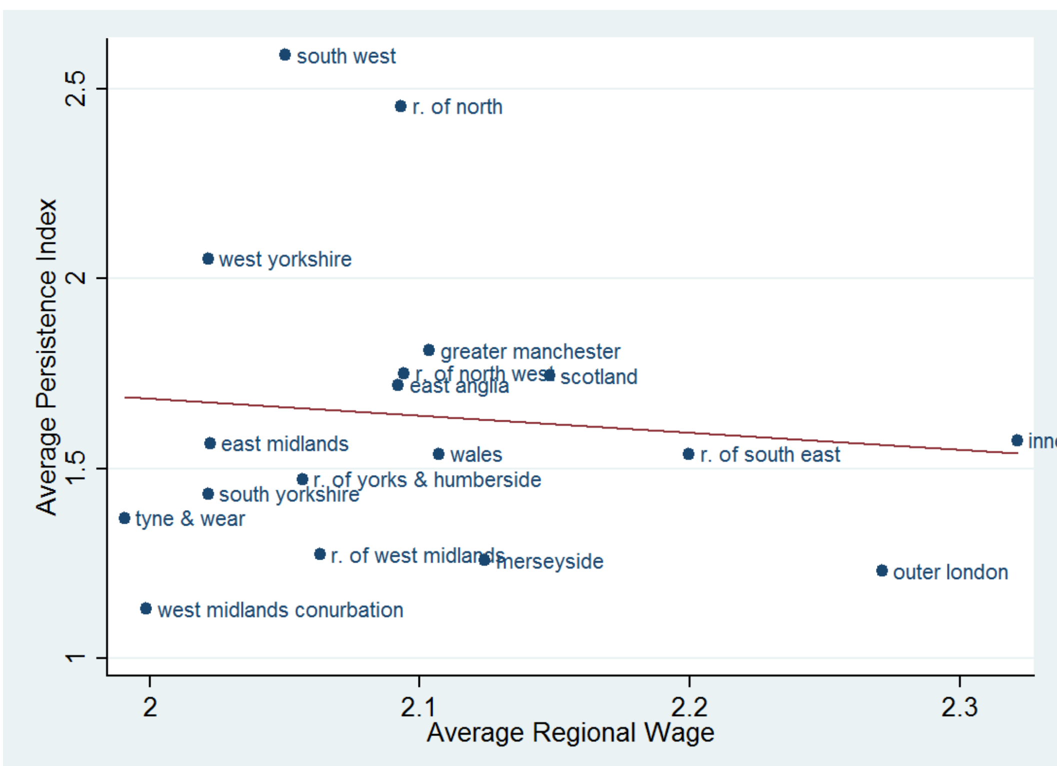

Figure 5. Plot of Average Persistence Index (weighted average of occupation–specific likelihood ratios, by region) vs. Average Regional Wage. Source: BHPS (1991–2008).

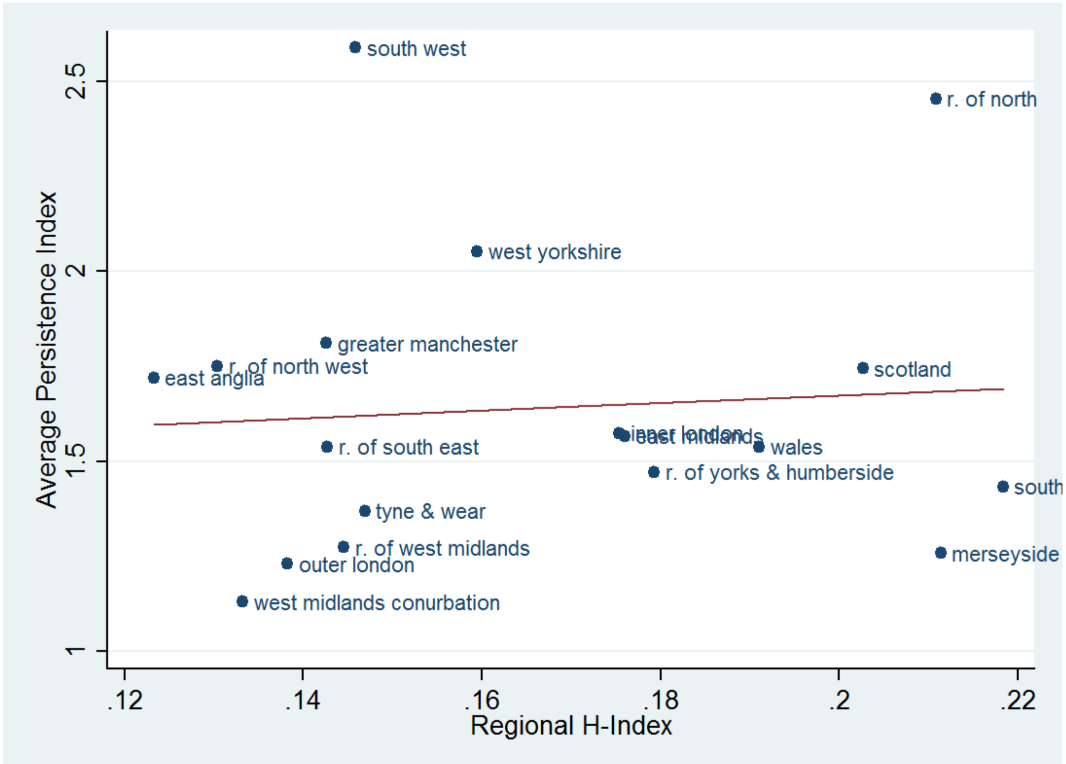

Figure 6. Plot of Average Persistence Index (weighted average of occupation–specific likelihood ratios, by region) vs. Herfindahl index of occupations (at the 1–digit level) . Source: BHPS (1991–2008).

Occupational sector (contemporaneous)	**Likelihood Ratio** **<20**	**20-25**	**25-30**	**30+**
Managers & Administrators	1.78	1.46	1.28	1.20
Professional	3.83	2.36	2.62	2.33
Associate Professional & Technical	2.05	0.91	1.86	2.27
Clerical & Secretarial	1.50	0.73	1.15	2.23
Craft & Related	1.32	1.48	1.73	1.68
Personal & Protective Service	1.19	1.53	2.69	0.61
Sales	1.55	1.45	0.98	1.49
Plant & Machine	1.27	2.07	2.06	1.89
Agriculture & Elementary	2.48	2.31	2.86	4.51
Average (unweighted)	1.88	1.59	1.91	2.03
Average (weighted)	1.62	1.49	1.83	1.92

Table XV. Occupational Persistence Indexes, by Son's Age Group (Likelihood Ratios). Source: BHPS (1991–2008).

Dependent Variable: Job–Finding Rate

	(1) POLS	(2) RE	(3) FE
Panel A			
Father in same occupation ($\pi_{i,t}$)	0.0481***	0.0497**	0.0564**
	(0.016)	(0.020)	(0.025)
Average in–sample JF	0.125	0.125	0.125
N	4098	4098	4098
R^2	0.055	-	0.047
Number of pairs	-	400	400
Panel B			
Father in same occupation ($\pi_{i,t}$)	0.0821***	0.0793**	0.0882**
	(0.025)	(0.032)	(0.039)
Average in–sample JF	0.115	0.115	0.115
N	2093	2093	2093
R^2	0.084	-	0.074
Number of pairs	-	212	212

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table XVI. Robustness Checks: Regressions of job–finding rate (transition from Unemployed to Employed); coefficient for **father in same occupation** (dummy variable), standard errors and average job–finding rate in the regression sample. Model 1 is a pooled OLS regression, model 2 is a random effects GLS regression, and model 3 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behavior, marital status, ethnicity, quarter and occupation of search/employment. In Panel A, we exclude the spells of self–employment from the estimation. In Panel B, we exclude all the workers who report having been self–employed at least once in their lifetime. Source: BHPS (1991–2008).

Dependent Variable: Job–Finding Rate

	(1)	(2)
Father in same occupation ($\pi_{i,t}$)	0.0404**	0.0437**
	(0.018)	(0.018)
Father with high tenure (dummy: 1 if above average) ($ht_{i,t}$)	-0.0266*	
	(0.014)	
Interaction term ($\pi^*_{i,t} ht_{i,t}$)	0.0634	
	(0.039)	
Log of father's tenure in years ($\log(t_{i,t})$)		-0.00512
		(0.006)
Interaction term ($\pi^*_{i,t} \log(t_{i,t})$)		0.0207
		(0.014)
N	4142	3726
R^2	0.059	0.062

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table XVII. Robustness Checks: Regressions of job–finding rate (transition from Unemployed to Employed); coefficient for **father in same occupation** (dummy variable), **father with high tenure** (dummy variable), **father's occ. tenure** (log), **interaction terms** and standard errors. Both models are pooled OLS regressions. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behavior, marital status, ethnicity, quarter and occupation of search/employment. Source: BHPS (1991–2008).

Dependent Variable: Log Hourly Wage					
	(1)	(2)	(3)	(4)	(5)
	POLS	RE	POLS	RE	FE
Panel A					
Share of time in occ. of father (q_i)	-0.134***	-0.155***			
	(0.016)	(0.031)			
Father in same occupation ($\pi_{i,t}$)			-0.073***	-0.032**	-0.007
			(0.013)	(0.013)	(0.014)
N	6324	6324	4664	4664	4664
R^2	0.617	-	0.602	-	0.639
Number of pairs	-	908	-	833	833
Panel B					
Share of time in occ. of father (q_i)	-0.0967***	-0.109***			
	(0.015)	(0.029)			
Father in same occupation ($\pi_{i,t}$)			-0.050***	-0.019	-0.003
			(0.012)	(0.012)	(0.012)
N	5664	5664	4159	4159	4159
R^2	0.547	-	0.540	-	0.606
Number of pairs	-	866	-	788	788

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table XVIII. Robustness Checks: Regressions of Log Hourly Wage; coefficient for **share of time spent in same occupation as father** (from 0 to 1), standard errors and **father in same occupation** (dummy variable). Models 1 and 3 are pooled OLS regressions, models 2 and 4 are random effects GLS regressions, and model 5 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education and occupation, second–order polynomials in occupational tenure and potential labor market experience, firm size, region of residence, smoking behavior, marital status, ethnicity and year. Panel A excludes from the estimating sample all wage observations above percentile 99 or below percentile 1. Panel B excludes from the estimating sample all wage observations above percentile 95 or below percentile 5. Source: BHPS (1991–2008).

Dependent Variable: Occupational Persistence Rate

	(1)	(2)	(3)	(4)
Father's log wage	0.0419***	0.0481***		
	(0.016)	(0.016)		
Log wage		-0.0779***		
		(0.017)		
Father's average log wage			0.0362*	0.0503***
			(0.019)	(0.019)
Average Log Wage				-0.0939***
				(0.020)
Average in–sample Persistence Rate	0.172	0.172	0.172	0.172
N	3467	3467	3467	3467
R^2	0.134	0.139	0.133	0.139

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table XIX. Regressions of Occupational Persistence (being in the same occupation as the father); coefficient of **father's log wage**, **log wage**, **father's average (lifetime) log wage**, **average (lifetime) log wage**, standard errors and average job–finding rate in the regression sample. All models are pooled OLS regressions. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behavior, marital status, ethnicity, quarter, and occupation of search/employment. Source: BHPS (1991–2008).

Dependent Variable: Occupational Persistence Rate

	(1)	(2)
Father's log wage above average	0.0231*	
	(0.012)	
Father's log wage above occ.–specific average		0.0267**
		(0.011)
Average in–sample Persistence Rate	0.165	0.165
N	4953	4953
R^2	0.135	0.136

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table XX. Regressions of Occupational Persistence (being in the same occupation as the father); coefficient of **father's log wage above average**, **father's log wage above occupation–specific average**, standard errors and average job–finding rate in the regression sample. All models are pooled OLS regressions. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behavior, marital status, ethnicity, quarter, and occupation of search/employment. Source: BHPS (1991–2008).

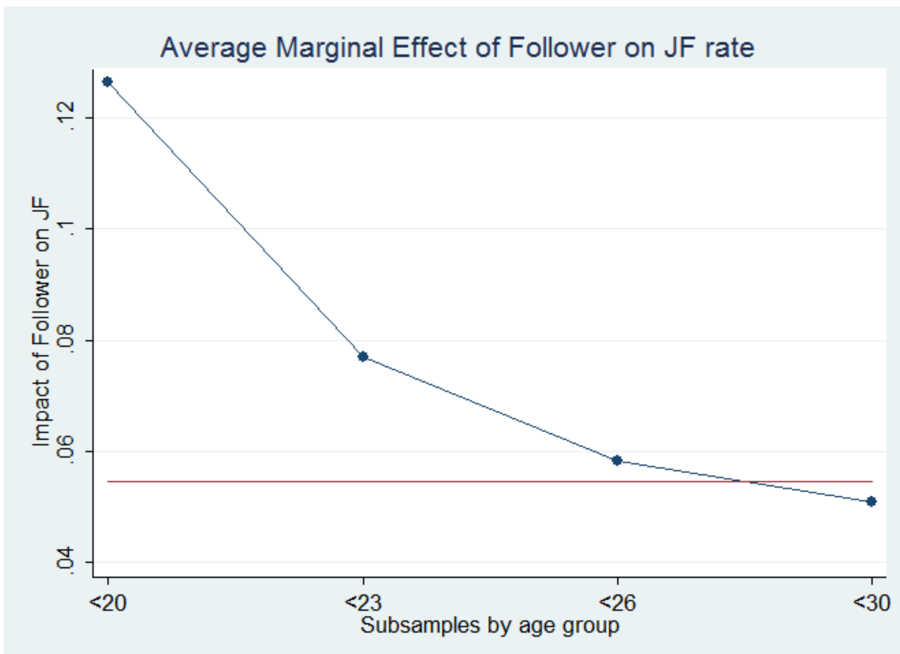

Figure 7. Average marginal effect of $\pi_{i,t}$ (coefficient of Column 1 of Table III), by age group. The red line is the average marginal effect for the entire sample. Source: BHPS (1991–2008).

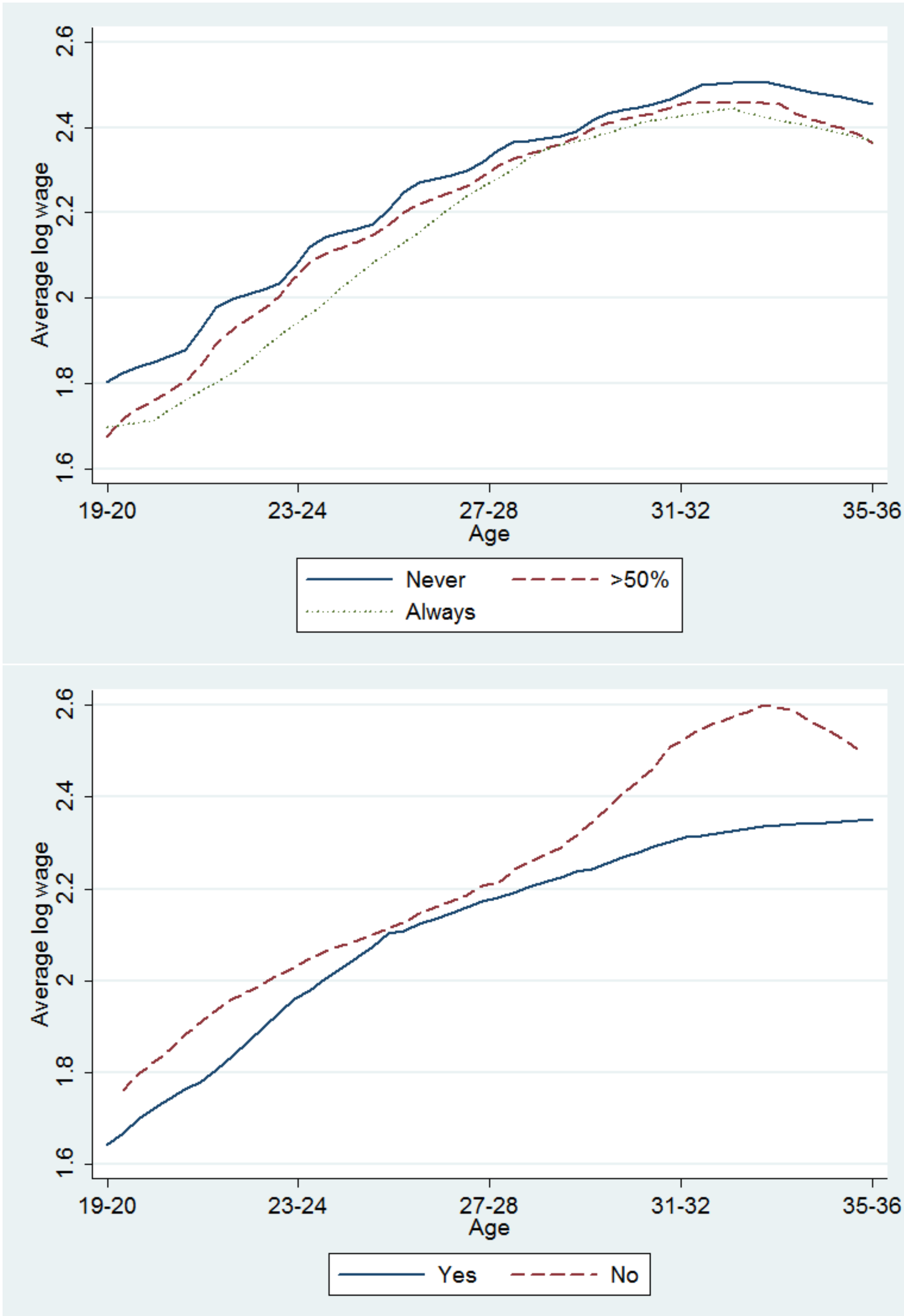

Figure 8. Wage profiles by proportion of job spell with the father in the same occupation. In the upper graph, groups are defined with respect to the entire job spell length. In the lower graph, groups are defined with respect to the start of the job spell. Source: BHPS (1991–2008).

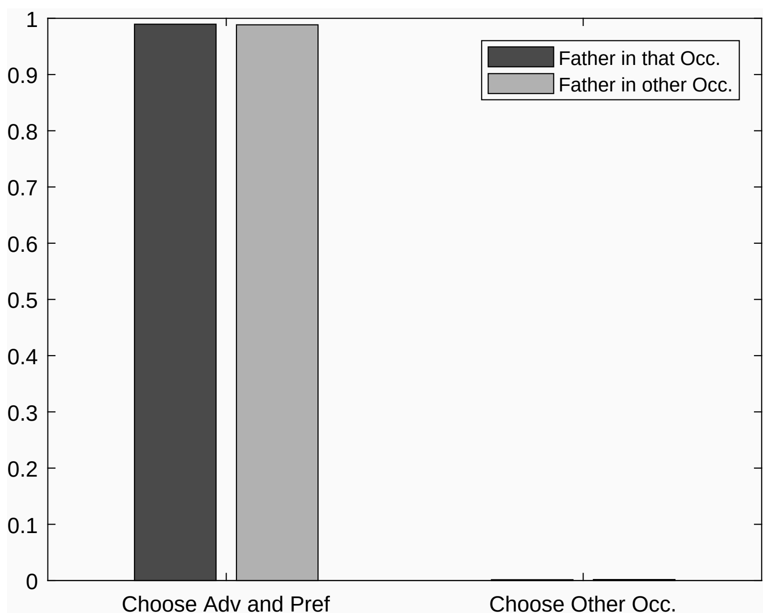

Figure 9. Probability of choosing occupations (Average Policy Function), for unemployed workers with comparative advantage and preference in the same occupation as the father. Source: BHPS (1991–2008).

Parameter	Description	Value	Target/Source	Data	Model
Intergenerational Transmission					
ξ	Transmission of Networks	0*	JF premium of followers (BHPS)	[0.055]	0.001
ρ_τ	Transmission of Comparative Advantage	1.000	Difference in Proportion of Followers by Father's wage (BHPS)	0.023	0.256
ρ_ϕ	Transmission of Preferences	0.111*	Interg. Occupational Persistence (Lik. Ratio, BHPS)	1.720	1.597
Heterogeneity and Laws of Motions					
$\hat{h}$	HC premium	0.260	Average occupational tenure returns after 5 years (BHPS)	-	-
$\hat{n}$	Networks premium	1.700	Proportion of jobs found through contacts (Pellizzari 2010)	0.230	0.229
$\hat{\tau}$	Comparative Advantage premium	1.042	Within–occupation log wage variance (BHPS)	0.166	0.181
$\bar{\phi}$	Baseline preference for jobs	0	Normalization	-	-
$\hat{\phi}$	Preference premium	0.882	Wage discount of followers (BHPS)	[0.076]	-0.316
p_h^+	Probability of accumulating HC (employed)	0.0166	Average occupational tenure returns after 5 years (BHPS)	0.260	0.260
p_h^-	Probability of losing HC (unemployed)	0.390	Average wage discount after unemp. (Arulampalam 2001)	0.076	0.072
p_n^+	Probability of accumulating networks (employed)	0.003	Regression of JF rate vs. past occupational tenure (BHPS)	0.008	0.008
p_n^-	Probability of losing networks (unemployed)	0.117	JF rate–unemployment duration profile (see text, BHPS)	1.066	1.094
σ	Standard deviation of preference shocks	0.348	Occupational change rate, after U spell (monthly, BHPS)	0.357	0.365
Environment					
O	Number of occupations	9	1-digit SOC aggregation	-	-
κ	Vacancy posting cost	4.765	Average UE rate (monthly, BHPS)	0.125	0.112
δ	Exogenous separation rate	0.003	Average EU rate (monthly, BHPS)	0.005	0.005
λ	Discount factor	0.9966	From literature	-	-
ζ	Age shock	0.00416	Average length of worklife: 20 (young) + 20 (old) years	40	40
b	Unemployment benefit	0.584	Average replacement rate (OECD)	0.530	0.509
χ	Surplus sharing rule	0.7	Normalization	-	-
A	TFP parameter of matching function	0.1	Normalization	-	-
γ	Elasticity of matching function w.r.t. unemp.	0.5	Normalization	-	-

Table XXI. Calibration Results, Restricted Model (* = restricted parameters; [] = non-targeted moments)

		($\kappa = 1.34$)				($\kappa = 15.00$)			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
No parental net.	($\xi = 0$)	-	✓	-	-	-	✓	-	-
No comp. adv. trans.	($\rho_\tau = 1/O$)	-	-	✓	-	-	-	✓	-
No pref. trans.	($\rho_\phi = 1/O$)	-	-	-	✓	-	-	-	✓
Occupational Persistence		1.254	1.248	1.005	1.254	1.683	1.143	1.579	1.586
(Δ% from baseline)		-	(-2.285)	(-98.112)	(-0.003)	-	(-79.136)	(-15.180)	(-14.250)
Welfare (Δ% from baseline)		-	(-0.324)	(-0.066)	(-0.000)	-	(0.094)	(-0.052)	(0.044)
Sorting along comparative advantage (sons)		0.999	0.999	0.999	0.999	0.604	0.612	0.599	0.609
Sorting along preferences (sons)		0.112	0.112	0.112	0.112	0.508	0.499	0.512	0.502
Sorting along comparative advantage (fathers)		0.924	0.924	0.924	0.924	0.584	0.588	0.581	0.587
Sorting along preferences (fathers)		0.187	0.187	0.187	0.187	0.528	0.524	0.531	0.525
Output per worker (Δ% from baseline)		2.364	(0.000)	(-0.001)	(0.000)	1.888	(0.549)	(-0.286)	(0.345)
Unemployment Rate (Δ% from baseline)		0.011	(1.866)	(0.307)	(0.001)	0.183	(-0.709)	(0.311)	(-0.279)

Table XXII. Occupational Persistence Decomposition and Welfare Analysis. Numbers in parentheses are relative changes from baseline. Counterfactual Experiments: Columns (1)-(4) are under the baseline calibration and the US level of labor market frictions; Columns (5)-(8) are under the baseline calibration and Spanish level of labor market frictions.