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

Salvatore Lo Bello* and Iacopo Morchio‡

November 2016

Job Market Paper - Click here for most updated version

Abstract

We develop a dynamic quantitative model of occupational choice and search frictions with multiple channels of intergenerational transmission (abilities, social contacts and preferences), and use it to decompose the occupational persistence observed in UK data. In our model, those who pursue the same occupation as their parents are more likely to base their decision on factors other than their productive advantage, generating misallocation in the economy. A key prediction of our theory is that occupational followers are expected to find jobs faster and earn lower wages, implications for which we find strong empirical support in UK data. We find that all three sources of persistence are quantitatively important: parental networks by themselves can account for 59% of total persistence, while ability and preferences transmission explain 25% and 23%, respectively. Shutting down parental networks or the transmission of preferences yields welfare gains (due to a better allocation of workers), while removing the transmission of comparative advantages generates a welfare loss. Moreover, we demonstrate that parental networks become less important in determining the allocation as search frictions vanish. Finally, increases in unemployment benefits reduce mismatch, thus having a positive effect on output per worker.

Keywords: Comparative Advantage, Labor Productivity, Mismatch, Occupational Mobility, Social Contacts.

JEL Classification Numbers: J24, J62, J64.

*Universidad Carlos III de Madrid. E-mail: sabello@eco.uc3m.es

†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, Betsy Caucutt, Juanjo Dolado, Javier Fernandez–Blanco, Manolis Galenianos, Manuel Garcia–Santana, Pedro Gomes, 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 participants at the PhD Student UC3M Workhop, 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) and Doctoral Workshop on Quantitative Dynamic Economics 2016 (Marseille) for useful comments and suggestions.

1 Introduction

It is well known that several economic outcomes are correlated across generations, most notably income, education and occupational choice.[1] 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, Güell et al. 2015). With respect to the occupational choice, a large degree of persistence might reflect the presence of barriers of various types in the labour markets, implying a suboptimal allocation of workers to jobs. However, a deep theory of occupational persistence, that would help us understand whether persistence is indeed associated to inefficiencies or not, has not been developed yet.

In this paper we study how intergenerational occupational persistence and labor misallocation are related. We demonstrate that when persistence is generated from several sources, it is crucial to measure the quantitative importance of each of the channels, and how they interact with one another. To do so, we develop a dynamic model of occupational choice and search frictions featuring multiple channels of intergenerational transmission, and use it to decompose the occupational persistence observed in UK data. Our model provides novel insights on how the different sources of persistence may impact 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 static 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 the occupation in which they currently work, which is not necessarily their offspring's comparative advantage. We also analytically establish that higher levels of persistence can be associated to higher or lower welfare, depending on the parametrization and whether they are the result of stronger transmission of productivity or social

[1] 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 education, see Chevalier *et al.* (2009). For occupational persistence, see Hout & Beller (2006), Constant & Zimmermann (2004), Escriche (2007), Eberharter (2008) and Dustmann (2004). Two reviews of the literature are in Black & Devereux (2010) and Ermisch *et al.* (2012).

contacts across generations. In our model, mismatch is not detrimental *per se* for welfare, as workers are in principle choosing optimally the occupation. However, search frictions interact with the level of mismatch present in the economy, an effect which is not internalized by workers. Thus, the equilibrium is generally inefficient because firms offer fewer jobs if they expect workers to be more mismatched and consequently less productive, as in Bentolila *et al.* (2010).

To assess the quantitative importance of the different channels, we then extend our theory to a dynamic model embedding a third channel of transmission (preferences) and allowing for mobility over the life–cycle. In a quantitative exercise, we find that all three forces are relevant in generating occupational persistence: social networks can account for about 59% of persistence, whereas transmission of comparative advantages and preferences account for 25% and 23%, respectively. Moreover, we demonstrate that a model in which only productive abilities are transmitted across generations falls short at accounting for several key pieces of evidence, and in general it provides a much worse fit to the data.

Our theory with multiple sources of transmission delivers two main predictions for agents who choose to pursue the same occupation as their parents (*occupational followers*), with respect to those who do not (*occupational movers*). First, the former will find jobs at a higher speed. Second, they will earn lower wages. The model also predicts that the first mechanism is a casual effect, while the second is due to a static selection mechanism and is expected not to survive fixed effects analysis. We find confirmation of all these 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 offspring, 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, workers are 72% more likely to be in a given occupation if their father is currently also employed in the same occupation. When we turn to testing the predictions of our model, we find that occupational followers experience finding rates that are 5–6 p.p. higher with respect to occupational movers. Given that the average in–sample job–finding rate is about 12.5 p.p., being in the same occupation as the 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 have large discounts (between 7 and 14%) with respect to occupational movers. Interestingly, the discount does not survive the inclusion of individual fixed effects. These results are robust to alternative definitions of the *occupational followers* group, using information on the entire labor market careers as well as only contemporaneous information.

We use our quantitative model to assess how each source of persistence affects welfare. When we shut down parental networks or the transmission of preferences, welfare improves, due to the better (more aligned with productive advantages) allocation of workers to occupations: output per worker goes up, and firms react posting more vacancies per worker than before. Instead, when we shut down the transmission of abilities, the reduction in persistence is accompanied by a reduction in welfare, driven by the worse allocation of the workforce. However, the changes in welfare are relatively small, of the magnitude of 0.1–0.2%.

We also investigate the role of search frictions, finding that the impact of parental networks on the allocation becomes negligible as frictions tend to zero; also, differentials in vacancy posting costs in our model generate at the same time differentials in unemployment rates and labor productivity. Finally, in our model more generous unemployment benefits imply that workers are less likely to choose the same occupation as their father to reduce the probability of facing unemployment, and therefore are less likely to be mismatched. This implies that increasing benefits can potentially be welfare–improving. Nonetheless, when the change in unemployment benefits is financed through lump–sum taxation, the overall effect on welfare is slightly negative. For instance, an increase in unemployment benefits of 25% yields an increase in output per worker of 0.4%, but a decrease in welfare of 0.1%. Importantly, the restricted model provides very different implications for this policy change, largely overestimating the negative welfare effects of increases in benefits, which become an order of magnitude larger.

While there is a significant amount of work studying why income is persistent across generations,

work on occupational choice is scarcer in the economic literature[2], to which we contribute along several dimensions. First, we add to the empirical literature studying occupational persistence across generations (Constant & Zimmermann (2004), Hellerstein & Morrill (2011), Ermisch & Francesconi (2002), Di Pietro & Urwin (2003), Long & Ferrie (2013)), by providing new estimates of likelihood of belonging to the same occupational category using *contemporaneous* information based on monthly transitions.

Second, our study bridges the literature that focuses on the determinants of occupational choice (Miller (1984); McCall (1991); Keane & Wolpin (1997); Papageorgiou (2013); Carrillo-Tudela & Visschers (2014)), its consequences for inequality (Kambourov & Manovskii (2009)) and unemployment duration (Wiczer (2014)) with the literature studying occupational persistence and career following (Laband & Lentz (1983)). Our paper is one of the few to take a theoretical and quantitative perspective on occupational choice across generations[3], providing novel insights on how persistence maps to efficiency and aggregate welfare.

Third, we relate to the literature studying social networks in the labor market (for instance, see Horváth (2014) and Galenianos (2014))[4] 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 having misallocation of the labor force due to socially suboptimal occupational choice, just as in Bentolila *et al.* (2010), Celik (2015), Hsieh *et al.* (2013) and Munshi & Rosenzweig (2016). Last, the decomposition exercise undertaken here is very close in spirit to other decomposition exercises of income persistence across generations, such as Restuccia & Urrutia (2004), Lee & Seshadri (2014), Abbott *et al.* (2013) and Gayle *et al.* (2015).

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

[3]In this sense, the closest paper to ours is the one by Sinha (2014), who studies how borrowing constraints affect occupational choices and how this mechanism can be important to understand persistence in developing countries.

[4]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).

The rest of the paper is organized as follows. Section 2 develops a static model of occupational choice and intergenerational transmission. In Section 3 we present our empirical evidence on occupational persistence and in support of our simple theory. Section 4 outlines the dynamic quantitative model, which is calibrated and used for counterfactual experiments in Section 5. Finally, we conclude in Section 6.

2 Model

This model is based on Bentolila *et al.* (2010), that is a static version of a standard search model á la DMP. With respect to their framework, we add the following features: i) overlapping generations connected to each other; ii) a notion of occupational persistence across generations; iii) two different sources of persistence across generations.

The economy is populated by a measure 2 of individuals who live for 2 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 (which we term "his father"). There are two occupations denoted by $j \in \{A, B\}$, each of which corresponds to a separate labor market. The main choice of individuals here 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 $M(U_j, V_j)$. We define the probability of meeting with a vacancy for an unemployed worker $p(\theta_j) = \frac{M(U_j,V_j)}{U_j}$ (per unit of search effort), where $\theta_j = \frac{V_j}{U_j}$ represents the labor market tightness. Conversely, $q(\theta_j) = \frac{M(U_j,V_j)}{V_j}$ represent the job-filling rate (probability of a firm to meet with a worker).

Upon matching, the surplus generated is split according to Generalized Nash Bargaining, where the worker's bargaining power is equal to β. We assume that search takes place during the young phase, while production only takes place during the old phase[5]: as a result, equilibrium wages are equal to a

[5]These assumptions are made for simplicity, as they keep the model static.

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, with $a > 0$. We assume that there exist two types of workers, with $\tau = \{A, B\}$ denoting the occupation in which one has his productive advantage. Types are transmitted across generations according to the following: $P(\tau = A|\tau^F = A) = P(\tau = B|\tau^F = B) = \rho$ (we denote 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 the following, we will assume $\rho > \frac{1}{2}$, that is individuals have a tendency to be of the same type as their fathers, on average.

Social contacts are occupation–specific and help individuals to find jobs, in that they increase 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 their 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 young to old, accumulate a network of size N with probability μ and of size n otherwise, with $\mu \in [0, 1]$. Summing up, we have that:

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

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

choices.

2.1 Equilibrium characterization

We focus on symmetric equilibria, given that the two markets have the same fundamentals[6]. In equilibrium, each market attracts a measure 1 of workers, with an identical composition of labor productivity and contacts endowment. In order to ease the exposition, we will sometimes explicitly look at occupation A, with the understanding that the same results hold for occupation B.

There is a number of workers in this economy for which the occupational choice is trivial. Specifically, all those workers for which the productive advantage lies in their father's occupation will simply choose that one as their occupation. We calculate their mass as follows[7]:

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

where o^F denotes the father's occupation. Noting that the 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 into 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 *mismatch* in this economy, that is how bad the sorting between workers and occupations is along the productivity dimension. If $m = 1$, the occupational choice is perfectly aligned with the type, that

[6] Thanks to the symmetry of the equilibrium, we will often 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).

[7] Throughout this model, we will often 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.

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 trade-off in their occupational decision.

In order to solve for the equilibrium m, let us first look at the optimal occupational choice of the workers who face the trade–off. These are the workers whose productive advantage lies in one occupation but whose father (with his social contacts) is in the other occupation. For instance, consider those for which $\tau = A$ 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 ($\beta(1+a)y$) more than compensate for the higher job–finding rate in occupation B ($p(\theta)(1+\tilde{n})$). It is easy to see that, under Assumption A, condition (3) holds true only for those workers endowed with a small network. All the other ones (whose mass is μ) decide optimally not to pursue their productive advantage.

Let us now look at the total probability that a 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 a type A chooses occupation A as a weighted average of him choosing occupation A depending on whether he has a trade–off or not. In the first case, he chooses A with probability 1; in the second case, with probability $(1-\mu)$. The expression in (2) has to be equal to 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. □

We have two extreme cases: $m^* = 1$ when $\mu = 0$ (nobody's choice depends on contacts) or $m^* = \frac{1}{2}$ when $\mu = 1$ (everybody follows their father's footprints, no matter what their type is). In general, m^* is decreasing in μ and increasing in ρ. The former effect is straightforward: the higher μ, the more workers choose purely according to contacts, whose availability need not be aligned to productivity. Instead, the reason for the latter 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) coincides more and more 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 the workers' sorting. As a result, the allocation of the workforce improves. In the limit case, when $\rho = 1$, the allocation of workers is not affected anymore by μ and there is no mismatch in the economy.

It is useful to derive an expression for the equilibrium value of Ψ, by replacing 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 trade–off 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 decide their occupation according to contacts' availability. In turn, this lowers Ψ, as more young workers will face a trade–off. Conversely, an increase in ρ improves the alignment of contacts and comparative advantage,

through an increase in m^*.

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

Define the total efficiency units of search as U:

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

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

Thus, a firm meets with a productive type with probability:

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

The expression above is in fact the relevant dimension of *mismatch* for aggregate outcomes, that is how bad the allocation of workers is perceived by firms. In general, it can be shown that $\gamma(m) < m$ for

$\mu > 0$. This is because mismatched workers are over–represented in the unemployment pool, displacing the better workers. It turns out that an increase in μ has a negative effect on $\gamma(m)$, for a twofold reason. Not only it alters the number of mismatched workers, but it also gives to them higher probability of being drawn by a firm. On the other hand, ρ has a positive effect on $\gamma(m)$: it raises the total efficiency units of search (through the effect on $\Psi(m)$, from equation 7), and it reduces the number of those with a trade–off (it raises $\Psi(m)$ itself).

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. □

Let us now look at the free–entry condition. The following equation equates expected profits of firms with the fixed cost of posting a vacancy.

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

It shows that equilibrium tightness closely depends on $\gamma(m)$, how hard it is to match with a productive type. This implies that whatever has a positive (negative) effect on $\gamma(m)$, 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 expected profits from a match and therefore reduces incentives to firms entry. As a result, equilibrium tightness goes down. 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.

Let us first derive an expression for the degree of occupational persistence across generations in this economy. Such an expression will be useful in the next calculations:

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

We denote as $\mathcal{P}(m)$ the degree of intergenerational occupational persistence, that is the probability that young workers choose the same occupation as their fathers[8]. Let us define π as an individual dummy that takes the value 1 if the worker is in the same occupation as their father, and 0 otherwise. Our model allows us to derive predictions on the following two groups: the *occupational followers* (for which $\pi = 1$) and the *occupational movers* (for which $\pi = 0$). Let us first consider the probability of finding a job. In this model, this is equal to the product between $p(\theta)$ and the efficiency units of search exerted, which we call s. Occupational followers can exploit their social contacts and therefore will have higher units of search than occupational movers.

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

The previous expression is derived noting that, among occupational followers, some workers will be mismatched (their mass is m^{π}, an endogenous object that we solve for in the following). These workers necessarily must have exploited large networks, and thus their efficiency units of search are $(1 + N)$

[8]It does not matter whether we define it in terms or joint or conditional probability, given that we are studying a symmetric equilibrium.

with probability 1; this is because otherwise they would have not given up their comparative advantage. On the other hand, the non–mismatched might have both large or small networks, respectively with probability μ and $(1-\mu)$. In any case, each of these subgroups has a total of efficiency units of search above 1. As a result, the job finding rate of occupational followers will be strictly higher than the one of occupational movers.

To be able to say something about wages, we first need to derive expressions for the probability of being a productive type, for each of these two groups. In other words, we need to solve for m^{π}:

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

(13)

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 $(1-m^{\pi})$ the *mismatch among occupational followers*. If $\mu > 0$, it turns out that m^{π} is strictly smaller than one. This means that 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 yields that there is no mismatch among occupational movers, i.e. $P(\tau = A | o = A; o^F = B) = 1$. This depends on the fact that, by construction, those workers who do not choose the same occupation as their father are not using contacts and therefore they necessarily need to have based their decision on their productive advantage.

Conditional on matching, the probability of meeting with a productive type among occupational followers reads as follows:

$$\gamma^{\pi}(m^{\pi}) = 1 - \frac{(1 - m^{\pi})(1 + N)\mathcal{P}(m)}{U^{\pi}(m^{\pi})}, \tag{14}$$

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

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

The previous equation consists of three terms: the first two represent the efficiency units of search exerted by the occupational followers who are not mismatched, who can be using a large or a small network; the last term, instead, represents the mismatched workers, who are necessarily using a large network.

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

$$E(w|\pi = 1) = \beta(1 + \gamma^{\pi}(m^{\pi})a)y < \beta(1+a)y = E(w|\pi = 0). \tag{15}$$

For $\mu > 0$, we have that $\gamma^{\pi}(m^{\pi}) < 1$, and therefore wages of occupational movers are, on average, strictly higher than 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 first effect is predicted to survive the inclusion of controls for the fixed productive type, while the second effect 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 enhances the level of available social contacts, and as a result the job–finding rate increases. This happens

regardless of the worker's type or occupation. On the other hand, the differences in wages are only driven 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 not affected anymore by the presence of the father.

Another key prediction of this model is that wages will be correlated across generations. In this economy, only two levels of wage are offered in equilibrium: $\beta(1+a)y$ to the non–mismatched workers, and βy to the mismatched workers. It turns out that having a mismatched father increases the probability for the offspring to be mismatched as well. Let us define $\bar{w}$ as a dummy that takes value 1 is the worker is non–mismatched (that is, he is a potential high–earner), and 0 otherwise. Using the symmetry of the equilibrium, we can write the following:

$$P(\bar{w}=1|\bar{w}^F=1) = P(\bar{w}=1|\bar{w}^F=1; o^F=A). \tag{16}$$

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, being a high–earner in occupation j). We can now rewrite the previous expression:

$$P(\bar{w}=1|\bar{w}_A^F=1) = P(\bar{w}_A=1|\bar{w}_A^F=1) + P(\bar{w}_B=1|\bar{w}_A^F=1), \tag{17}$$

using the fact that $\bar{w}_A=1$ and $\bar{w}_B=1$ are mutually exclusive events. The first term of Equation (17) can be calculated as follows:

$$P(\bar{w}_A=1|\bar{w}_A^F=1) = P(o=A; \tau=A|o^F=A; \tau^F=A) = \tag{18}$$

$$= P(o=A|\tau=A; o^F=A; \tau^F=A)P(\tau=A|o^F=A; \tau^F=A) = \rho. \tag{19}$$

Instead, the second term of Equation (17) reads as follows:

$$P(\bar{w}_B = 1 | \bar{w}_A^F = 1) = P(o = B; \tau = B | o^F = A; \tau^F = A) = \tag{20}$$

$$= P(o = B | \tau = B; o^F = A; \tau^F = A) P(\tau = B | o^F = A; \tau^F = A) = (1 - \mu)(1 - \rho). \tag{21}$$

Substituting the last two equations into (17), we get:

$$P(\bar{w} = 1 | \bar{w}^F = 1) = \rho + (1 - \mu)(1 - \rho). \tag{22}$$

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

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

1. $P(\bar{w} = 1 | \bar{w}^F = 1) > P(\bar{w} = 0 | \bar{w}^F = 1)$.

2. $P(\bar{w} = 0 | \bar{w}^F = 0) > P(\bar{w} = 1 | \bar{w}^F = 0)$.

Remarkably, this result does not hinge on any transmission of efficiency level (i.e., unidimensional productivity) across generations, as it is usually done in the literature. In our case, the transmission of wage across generations is a by–product of the *transmission of mismatch* across generations.

2.3 Occupational Persistence, Mismatch and Efficiency

In this section we study how changes in the structural parameters of this economy (ρ and μ) affect the occupational persistence across generations. Moreover, we perform welfare analysis and investigate how

occupational persistence relates to welfare.

In this economy, occupational persistence is brought about by two different channels: use of contacts and transmission of type (comparative advantage). Strengthening any of these two channels (by increasing μ or ρ, respectively) has a positive effect on persistence.

Proposition 1. *The probability that young workers choose the same occupation as their father's, $\mathcal{P}^*$, is strictly increasing in both μ and ρ.*

Proof. See Appendix. □

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

$$\mathcal{P}(m^*) = \mu + (1-\mu)\Psi(m^*). \tag{23}$$

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 trade–off. These two effects go in different directions, but it turns out that the former effect is always stronger than the latter.

Instead, an increase in ρ affects the persistence only through changes in Ψ^*. In particular, higher levels of transmission of type improve the overall allocation of workers and therefore reduces the probability of facing a trade–off. As a result, occupational persistence also increases.

We now look more in depth at the relationship between occupational persistence, mismatch and welfare. As a first thing, we derive an equation relating the overall level of mismatch of the economy to

the mismatch among the occupational followers and the degree of occupational persistence:

$$\begin{aligned} m =& P(\tau = A|o = A; o^F = A)\mathcal{P}(m) + P(\tau = A|o = A; o^F = A)(1 - \mathcal{P}(m)) = \\ =& m^{\pi}\mathcal{P}(m) + (1 - \mathcal{P}(m)). \end{aligned} \tag{24}$$

That is, overall mismatch can be rewritten as a weighted average of the mismatch among occupational followers and among occupational movers. Plugging equation (10) into (24) and evaluating it in equilibrium yields:

$$m^* = \Psi(m^*) + (1 - \mathcal{P}(m^*)). \tag{25}$$

An implication of (25) is that $(1-m^*) \leq \mathcal{P}(m^*)$, i.e. the degree of mismatch of this economy is bounded above by the degree of occupational persistence. This is a reflection of the fact that only occupational followers happen to be mismatched[9].

We now turn to efficiency, and try to establish the relationship between occupational persistence and welfare. Let us define welfare as aggregate net income, that is the difference between expected match output and the vacancy cost, multiplied by the measure of unemployed in the economy.

$$\mathcal{W} = 2[p(\theta)(1 + \gamma a)y - \kappa\theta]U. \tag{26}$$

It is easy to show that the equilibrium is generally inefficient, unless we are in a knife–edge case. In particular, the derivative of welfare with respect to market tightness reads as follows:

$$\frac{\partial \mathcal{W}}{\partial \theta} = 2k \left[\frac{\eta_V(\theta) - (1 - \beta)}{1 - \beta} \right] U. \tag{27}$$

[9]This depends on the fact that if a worker is mismatched, it is because he is using social contacts, and in turn 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 not be the case anymore. For instance, we could allow workers to have social contacts in the other occupation from 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.

where we have used the fact that $q(\theta)(1+\gamma a)y = \frac{\kappa}{1-\beta}$ and the definition $p(\theta) = q(\theta)\theta$, with $\eta_V(\theta) = \frac{q'(\theta)\theta+q(\theta)}{q(\theta)}$ being the elasticity of the matching function with respect to V. Generally, $\eta_V(\theta)$ will be different from $(1-\beta)$, making the equilibrium inefficient[10].

Changes in μ or ρ have, in principle, an ambiguous effect on welfare. More specifically, both of these parameters affect at the same time three different things: 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 reads as follows:

$$\frac{d\mathcal{W}}{d\mu} = \frac{\partial \mathcal{W}}{\partial \mu} + \frac{\partial \mathcal{W}}{\partial \theta}\frac{d\theta}{d\mu}. \tag{28}$$

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

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

$$\frac{\partial \mathcal{W}}{\partial \mu} = 2\left[p(\theta)(1+\gamma a)y - \kappa\theta\right]\frac{\partial U}{\partial \mu} + 2p(\theta)ayU\frac{\partial \gamma}{\partial \mu}. \tag{29}$$

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.

[10] This economy has 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.

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

$$\frac{d\mathcal{W}}{d\rho} = \frac{\partial \mathcal{W}}{\partial \mu} + \frac{\partial \mathcal{W}}{\partial \theta}\frac{d\theta}{d\rho}. \tag{30}$$

The same considerations made above apply here. One important difference here is that the effect of ρ on market tightness is positive ($\frac{\partial \theta}{\partial \rho} > 0$). Moreover, the direct effect (the first term) is now clearly positive:

$$\frac{\partial \mathcal{W}}{\partial \rho} = 2\left[p(\theta)(1+\gamma a)y - \kappa\theta\right]\frac{\partial U}{\partial \rho} + 2p(\theta)ayU\frac{\partial \gamma}{\partial \rho}. \tag{31}$$

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

An implication of what we have discussed so far is that, in general, there is no one–to–one relationship between occupational persistence and welfare.

Proposition 2. *Changes in the degree of occupational persistence, $\mathcal{P}$, can be associated with both increases or decreases 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 Appendix. □

In words, it is crucial to understand what is the channel that is 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 to occupations. If, instead, persistence is brought about by workers choosing

according to their father's contacts' availability, then it becomes symptomatic of an underlying mismatch.

In this sense, unless we are able to decompose the sources of occupational persistence across generations, it is unclear whether the latter is socially desirable or not.

In Section 4, we will construct a structural model of intergenerational transmission and occupational choice, so that we will be able to perform the decomposition exercise in Section 5.

3 Empirical Evidence

In this section we document the degree of occupational persistence across generations, as well as 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 we use the same dataset constructed by Lo Lo Bello & Morchio (2013).

3.1 The Data

The BHPS is a yearly survey taken by about 10,000 individuals per year in the UK. It started in 1991, and the last available wave for this study is 2008. The survey is characterized by a quite high follow–up rate, with more than 90% of the individuals being interviewed also in the subsequent year. Besides this, each year a number of new households enters the sample. In total, 32,377 individuals were interviewed in the BHPS in the period 1991-2008. We restrict our sample to males aged between 16 and 65, and eventually we are left with 12,982 individuals, for a total of 1,023,888 monthly observations. Individuals report every year a detailed job history of the previous year, listing all the employment/unemployment spells, along with several characteristics of the job (among which, the occupational group). In this way, we can construct long time series for each individual (potentially up to 216 months) and, more importantly, we can identify transitions at the monthly frequency. Along with a detailed job history, for each individual a large amount of information is typically available, including sex, age, education,

occupation, race, marital status, region of residence and much more. One key feature of this dataset is that it allows us to connect individuals to their fathers and to follow them together over time.
The job–finding rate is defined as the probability of transiting from unemployment to employment, at the monthly frequency. Wages are calculated 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 recorder yearly).

3.2 Intergenerational Occupational Persistence

Our data allow us to study the extent of occupational persistence across generations. We compute the occupational distribution of workers and study the probability with which father and offspring belong 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 our *persistence index* $\mathcal{P}_j$ as the ratio between the probability of belonging to a given occupation j conditional on the father being also affiliated to the same occupation 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 1.

We find a large degree of occupational persistence. Our estimates indicate that likelihood ratios of occupational persistence are always above 1, indicating that workers are more likely to belong to a given occupation if their father is also in that occupation rather than somewhere else. The average likelihood ratio is estimated to be 1.76, meaning that individuals have 76% more chances to be in a

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 1. Occupational Persistence Indexes (Likelihood Ratios).

given occupation if their father belongs to that. This ranges between 29% and 167%, depending on the occupation. Interestingly, we do not find a specific trend of persistence across occupations (that is, persistence is neither higher or lower in high–paying or high–skill occupations)[11]. Repeating the same exercise at the 2–digit level, we find that the average unweighted likelihood ratio is 5.69, and the weighted one is 4.71 (see Table 14 in the Appendix).

Part of this persistence might be explained by usual socioeconomic variables such as age, education or the region of residence. In order to account for this, we estimate linear probability models in which we regress the probability of belonging to a given occupation (against any other) adding a number of covariates. We estimate such a model for each occupation, with the results shown in Table 2. The estimates reveal that workers are between 1.59 and 15.1 p.p. more likely to belong to the same occupation as their father, *coeteris paribus*. These are large probability differences, highly statistically significant in each of the 9 occupations.

[11] This seems to suggest that borrowing constraints are not playing a large role in the occupational choice. We also checked whether, within an occupational group, likelihood ratios vary by father's income. Our results show that occupational persistence decreases only slightly. For instance, if we only consider the top 1/3 earners in each of the occupations, the average likelihood ratio is 1.61 (Table 15 in the Appendix).

	(1) Occ. 1	(2) Occ. 2	(3) Occ. 3	(4) Occ. 4	(5) Occ. 5	(6) Occ. 6	(7) Occ. 7	(8) Occ. 8	(9) Occ. 9
Father in occupation 1	0.0335***								
	(0.003)								
Father in occupation 2		0.0365***							
		(0.003)							
Father in occupation 3			0.0550***						
			(0.005)						
Father in occupation 4				0.0411***					
				(0.005)					
Father in occupation 5					0.151***				
					(0.004)				
Father in occupation 6						0.0398***			
						(0.005)			
Father in occupation 7							0.0159***		
							(0.006)		
Father in occupation 8								0.127***	
								(0.003)	
Father in occupation 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 2. Regressions of Occupational Choice (dummy that takes value 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, quarter.

We also perform some more checks in order to investigate whether occupational persistence is mostly related to regional factors. For instance, belonging to a poor region (or a region where the variety of jobs supplied in the market is limited) might mechanically increase the likelihood ratio. In this sense, 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 the Herfindahl index of occupations, and find that neither of these two dimensions can predict persistence (results are shown in the Appendix, Figures 4 and 5). This is reassuring that regional factors are not playing too much of a role in our results[12].

3.3 Occupational Persistence and Occupational Attachment

So far we have shown that workers' occupational affiliation tends to be highly correlated with the one of their fathers. In this section we investigate whether the phenomenon of occupational persistence is persistent over the life–cycle. It is important to study this, as young workers –who are potentially sampling different occupations– might be the main responsible behind the likelihood ratios estimated in the previous section. More importantly, these workers might be using the father's occupation in a *stepping–stone* fashion (possibly to avoid unemployment), for then rapidly moving on with their career. If this was the case, then occupational persistence would mainly be a short–run phenomenon, with 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 16 in the Appendix).

Second, we study the length of the occupational spells of followers, as opposed to the ones of movers. The average occupational tenure is 2.16 years for the followers, and 1.73 years for the movers[13]. This is

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

[13] When we restrict our attention only to the spells that we observe from the start, we find again that followers tend to

true also for the occupations chosen very early in the career. Figure 1 plots the share of workers still in the first occupation against the years of labor market experience, separately for followers and movers (this latter variable is here defined at the moment of the start of the spell).

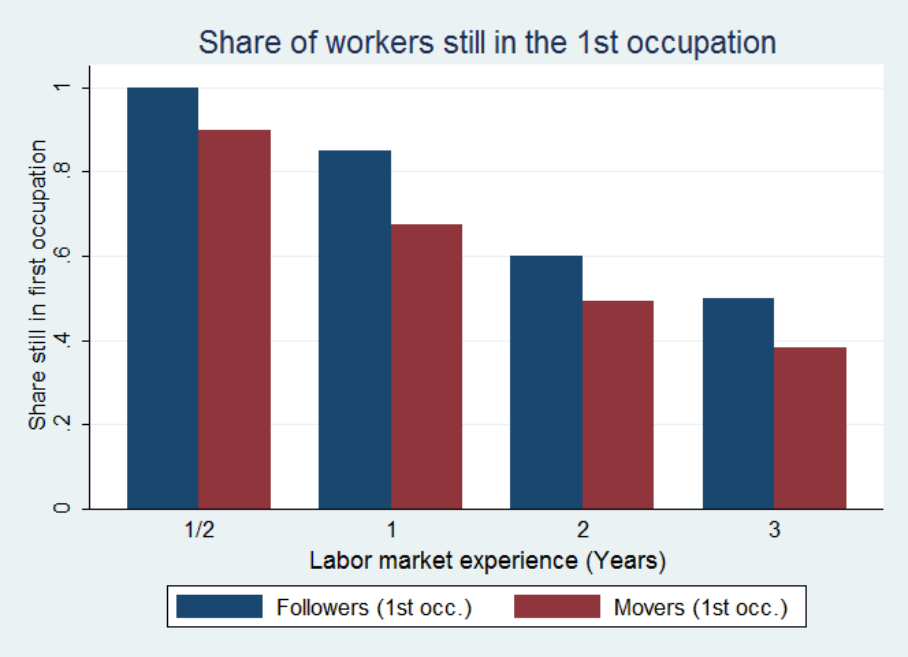

Figure 1. Share of workers still in the first occupation, by years of labor market experience. Source: BHPS.

We can see that those workers who start their career in the same occupation as their father's are substantially less likely to experience occupational mobility. For instance, after 2 years from the start of their first employment spell, 60% of occupational followers will not have experienced yet an occupational change[14], against 49% of occupational movers. At the same time, these statistics reveal that a large degree of *hysteresis* characterizes the occupational transitions. In other words, the initial occupation is a good predictor of the occupational affiliation even years after the start of the spell. In this sense, the father's influence on the initial occupational choice may have long–lasting consequences for the workers' outcomes and the aggregate allocation.

As an additional piece of evidence, we study whether the contemporaneous presence of the father

be more attached to their occupation (average tenure is 1.84 years against 1.69).

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

in the same occupation has an impact on the odds of changing occupations. We run the following regression:

$$OC_{i,t} = \alpha + \beta\pi_{i,t-1} + \boldsymbol{\gamma}\boldsymbol{X}_{i,t} + \epsilon_{i,t}, \tag{32}$$

where: $OC_{i,t}$ is a dummy taking the value 1 if the occupation at time t is different from the one at $t-1$ (there is an occupational switch[15]) and 0 otherwise; $\pi_{i,t}$ is a dummy variables that takes the value 1 if the occupation of offspring 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 the health level), region of residence and quarter dummies; $\epsilon_{i,t}$ is the idiosyncratic error term.

We estimate Equation 33 with Pooled OLS, Random Effects and Fixed Effects, with the estimates of β shown in Table 3.

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 3. 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, 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, occupation of employment.

We find that having the father employed in the same occupation is associated with a substantial

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

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 occupational change rate at the monthly frequency (2.65 p.p.). One possible interpretation could be 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, 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), making this interpretation less likely; ii) in Column 3 we are controlling for individual fixed effect, definitely ruling out this kind of explanations. The estimated coefficient –quite stable across specifications– suggests indeed that workers are less reluctant to switch occupation when this coincides with the one of the father, even on top of any unobserved fixed heterogeneity.

In the following subsections, we turn to testing the two main predictions of our model, namely that occupational followers[16] have higher job–finding rates but lower wages, on average.

3.4 Occupational Persistence and Job Finding Rate

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

$$JF_{i,t} = \alpha + \beta\pi_{i,t} + \boldsymbol{\gamma}\boldsymbol{X}_{i,t} + \epsilon_{i,t}, \tag{33}$$

where: $JF_{i,t}$ is defined only for the unemployed, taking the value 1 if a job is found at time t and 0 otherwise; $\pi_{i,t}$ is a dummy variables that takes the value 1 if the occupation of offspring i and his father coincide at time t[17], and 0 otherwise; $\boldsymbol{X}_{i,t}$ is a vector of control variables that include a third degree age

[16] Our working definition of *occupational follower* is based on $\pi_{i,t}$. However, what we document in the following 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).

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

polynomial, dummies for educational categories and occupational groups (observed for the employed, imputed for the unemployed), marital status, ethnic group, smoking behavior (to capture the health level), region of residence and quarter dummies; $\epsilon_{i,t}$ is the idiosyncratic error term.

We estimate Equation 33 with Pooled OLS, Random Effects and Fixed Effects, with the estimates of β shown in Table 4.

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 4. 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, 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, occupation of search/employment.

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%, having the father in the same occupation increases the chances of finding employment by about 42%. Importantly, the effect is robust to the inclusion of individual fixed effects (column 3), that control for unobserved heterogeneity across individuals. The identification of the individual fixed effects is allowed by the panel structure of our data. The coefficient in column (3) of Table 4 is estimated by exploiting variation in π (the presence of the father in the same occupation or not), *within* the offspring's worklife. This is in line with our model, that predicts that even after controlling for the occupation and the fixed type (τ), the father's occupational affiliation is still an

important determinant of the individual job finding rate.

We also find that these effects are robust to the exclusion of the self–employed from the estimating sample (Table 17, in the Appendix).

Finally, we study 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 6 in the Appendix). This piece of evidence lends support to our interpretation: young workers, who lack experience in the market, are the ones who are expected to depend more heavily on their father's contacts. In a dynamic setting, such as the one that we will develop in Section 4, workers accumulate contacts themselves, and therefore the influence of their father will fade out over time.

3.5 Occupational Persistence and Wages

The second key prediction of our model is that *occupational followers* will have on average lower wages. A share of those who choose the same occupation as their father's do not pursue their comparative advantage, and therefore they will have on average lower wages. However, according to our model this is entirely due to selection. In fact, upon choosing an occupation, the wage only depends on the individual productivity level in that occupation. That is, the father's occupational affiliation does not matter anymore. Thus, the difference between *occupational followers* and *occupational movers* is predicted to disappear after controlling for individual fixed effects (that capture the productivity level).

In order to test for this hypothesis, we construct for each individual an index of how much time he spent in the same occupation as his father, over his worklife. We do this for two reasons: first, we want to construct groups that do not change over the life–cycle of individuals; 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[18]. Thus, we define q_i as the fraction of employed worklife spent by individual i in the

[18]In the appendix (Figure 7) we show that differences in wages are robust to alternative definitions of *occupational followers*.

same occupation as his father:

$$q_i = \frac{\sum_t \pi_{i,t}}{\sum_t E_{i,t}}, \tag{34}$$

where $E_{i,t}$ is a dummy taking value 1 if the individual is employed in period t and 0 if he is unemployed. The index q_i ranges from 0 to 1, and measure for how many months (out of all those that an individual has spent employed) his occupation coincided with the one of this father. In Figure 2, we plot the wage profiles of three groups: those with $q_i = 0$, $q_i >= 0.5$ and $q_i = 1$.

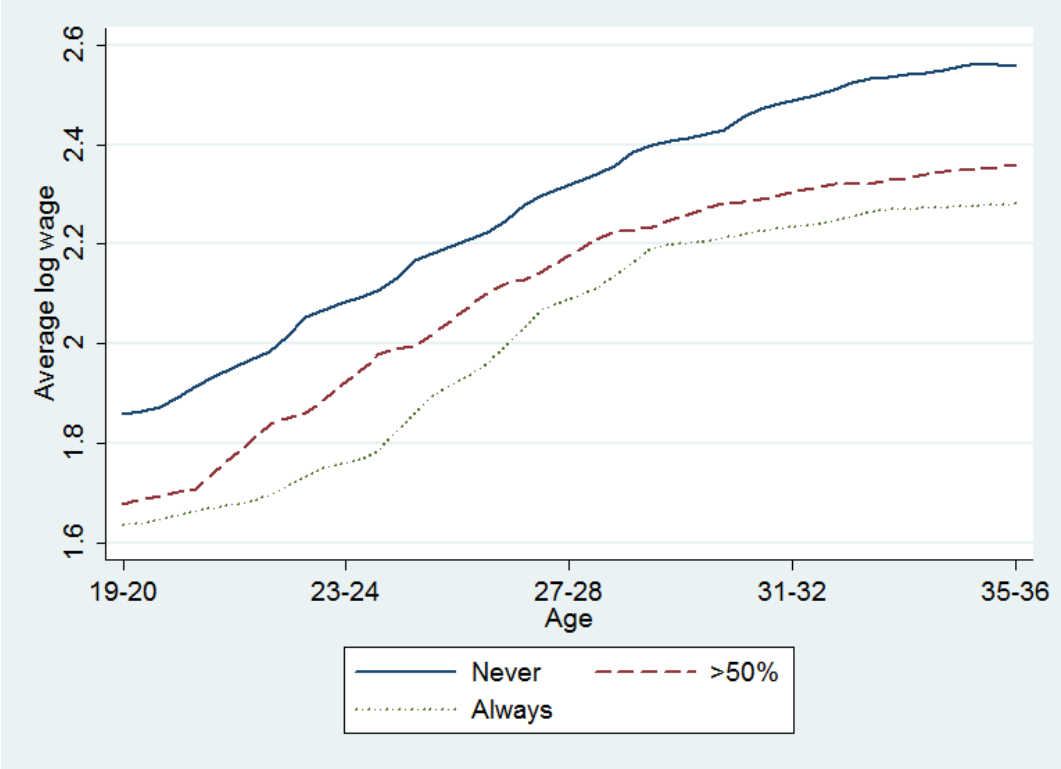

Figure 2. Wage profiles by proportion of employed worklife spent in the same occupation as the father. Source: BHPS.

We can see that those who spent more time in the same occupation as their father experience wages that are on average lower by up to 20%, throughout their entire life. It is remarkable how 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:

$$\log(w_{i,t}) = \alpha + \beta q_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \epsilon_{i,t}, \tag{35}$$

where: $\log(w_{i,t})$ is the natural logarithm of the wage (observed at the yearly frequency); q_i is defined in (34); $\boldsymbol{X}_{i,t}$ is a vector of control variables that include a third degree age polynomial, dummies for educational categories and occupational groups, a second order degree occupational tenure and potential labor market experience polynomial, marital status, ethnic group, smoking behavior (to capture the health level), firm size (proxy for firm productivity), region of residence and year dummies; $\epsilon_{i,t}$ is the idiosyncratic error term.

We estimate Equation 35 by POLS and RE. Given the impossibility of controlling for unobserved heterogeneity in this regression model, we estimate again the same regression substituting the independent variable q_i with $\pi_{i,t}$, the time–specific persistency dummy variable:

$$\log(w_{i,t}) = \alpha + \beta \pi_{i,t} + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \epsilon_{i,t}, \tag{36}$$

The first two columns of Table 5 show tell us that individuals who have spent more time in the same occupation as their father tend to earn lower wages, even after adding a number of controls. The discount associated with being in the same occupation is estimated to be about 14.5-15.2%, large and strongly statistically significant. Columns 3 to 5 of the same table shows estimates of β in Equation (36). We find an average discount of 7.6% associated to the presence of the father in the same occupation. However, this effect decreases to 2.4% (barely statistically significant at the 90% confidence level) when we allow for RE in the regression model, and it decreases further (it is neither statistically nor economically significant) when including individual fixed effects in the regression. Overall, these results lend strong support to our theory.

In the Appendix (Table 18), we show that these results are robust to trimming our sample (cutting top and bottom 1% and 5% of our 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. ($\pi_{i,t}$)			-0.0763***	-0.0239*	-0.000290
			(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 5. 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, model 5 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education and occupation, a second order degree occupational tenure and potential labor market experience polynomial, firm size, region of residence, smoking behavior, marital status, ethnicity, year.

3.6 Unemployment Risk and Wages

So far, we have established that occupational followers: i) among the unemployed, tend to spend less time in unemployment; ii) among the employed, tend to earn lower wages. However suggestive, these two pieces of evidence *per se* do not imply that individuals actually face the trade–off (between employment prospects and wages) that we depicted in our model.

Hence, here we try to produce some more compelling evidence, exploiting the entire career of the workers in our 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 their employment prospects) and the average monthly wage[19] earned throughout his career $\bar{W}_i$ (a measure of lifetime labor earnings). In order to compute these lifetime statistics, we consider observations from age 25 onwards[20].

[19] This measures incorporates the unemployment risk margin, too. We construct it in the following way: first, we multiply the monthly wage by the number of months that the individual is employed, for each year; then, we sum all of them up over the years, and finally we divide the total by the number of wage observations (to correct for the unbalancedness of our panel).

[20] The rationale behind this choice is that we want to make sure not be capturing things that have to do with different entry ages in the labor market.

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 have at the same time better employment prospects and lower wages.

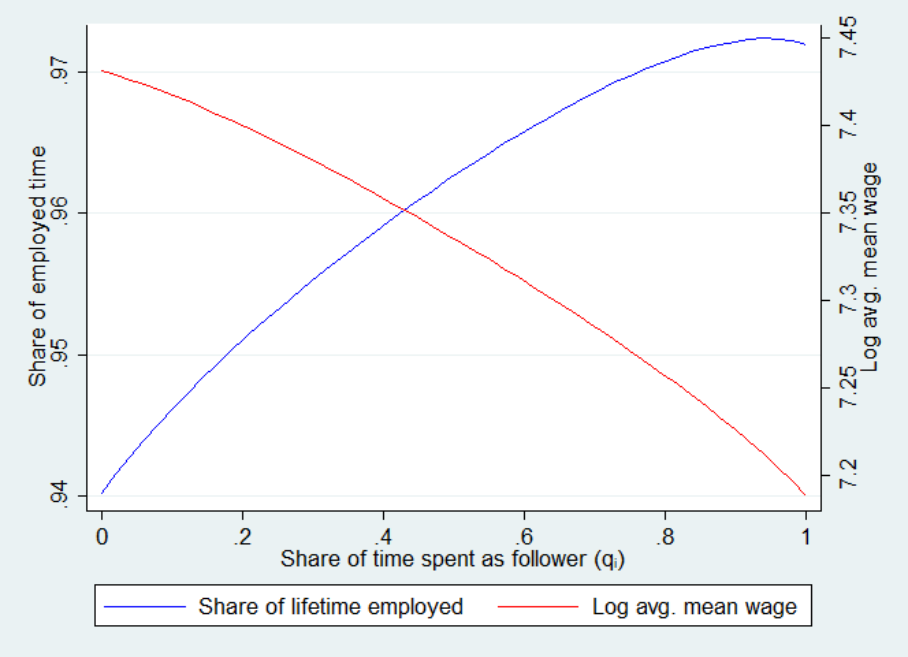

Figure 3. Locally weighted linear polynomial regression (degree 1, bandwidth 0.5) of share of lifetime employed and log average mean wage against the share of time spent as a follower. Source: BHPS.

In Table 6, we show the residual partial correlation between the aforementioned variables, controlling for fixed characteristics of individuals (education and race). Column 1 and 2 are a simple reflection of Figure 3: occupational followers tend to have lower wages (by up to -24%) but better employment prospects (on average, they spend as employed 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 correlation with q_i has different signs. The sign of both of these correlations is robust to the introduction of the other variables as a control. That is, conditional on lifetime employment prospects, followers tend to have lower wages (Column 4); and conditional on the average lifetime wage, followers tend to spend more time employed (Column 5).

Taken as a whole, the empirical evidence presented here directly speaks to the tradeoff predicated in our simple model, strongly suggesting that workers face such a tradeoff indeed.

In the next sections, we build a dynamic quantitative model of occupational choice, through which

	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 6. Regressions of Log average mean wage ($\bar{W}_i$) and Share of lifetime 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 employed; standard errors in parenthesis. All models include dummies for education and race.

we perform the decomposition exercise and study the welfare implications of occupational persistence.

4 Quantitative Model

In this section we develop a quantitative model that extends the previous model in a number of dimensions. First, the model presented here is dynamic, with individuals facing a stochastic ageing process. Second, we have O different occupations. Third, we introduce non–pecuniary benefits (preferences) for occupations, which are allowed to be correlated across generations. Fourth, we introduce occupation–specific human capital. We retain the assumption that social contacts are occupation–specific, and that fathers help their offspring to find jobs in their occupations, without internalizing that in their own occupational choice. Both social contacts and occupation–specific human capital evolve over time. Finally, we also introduce shocks to non–pecuniary benefits, in order to generate occupational mobility over the life–cycle.

4.1 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 an 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[21].

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 $h_{o,t}^i$ and social capital $n_{o,t}^i$ (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

[21] This assumption is made for simplicity and does not alter our results.

benefits associated to 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 $\boldsymbol{\phi}^{i,T}_t$, as well as the comparative advantage τ^i, are drawn at birth (i.e. at the entry in the labor market) and do not change over time, sticking to the worker for the rest of his life. Instead, the temporary component of preferences as well as occupation–specific human and social capital, evolve over time according to law of motions to be specified in the following. Each worker is either employed or unemployed ($e^i_t \in \{0, 1\}$), and is attached to some occupation o^i. Unemployed workers receive an unemployment benefit equal to b per period.

Let us denote parental variables with an F superscript, so that the occupation of individual i's father will be denoted $o^{i,F}$, parental networks $n^{i,F}_o$ and so on. Employed parents help their unemployed offspring by passing vacancies of which they become aware on to them. We assume that unemployed fathers do not help their offspring, as they are actively seeking a job for themselves.

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

4.2 Search and Reallocation across Occupations

We assume that search is costless and directed across occupations. Workers start their worklife as unemployed, and have to decide in which occupation to look for vacancies. The lack of firm heterogeneity usually implies that there is no scope for on–the–job search. Nonetheless, in our model employed parents do engage in on–the–job search, in order to help their unemployed offspring to find a job. Whenever they find a vacancy in their occupation, they will pass it on to their offspring. As a consequence, unemployed offspring find vacancies in their father's occupation with higher probability than anywhere else, *ceteris paribus*.

Each occupation consists of a single labor market, where the number of matches between unemployed

workers and vacancies is governed by the following constant returns to scale technology:

$$M_{o,t} = (U_{o,t})^{\eta}(V_{o,t})^{1-\eta} \tag{37}$$

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 offspring are currently unemployed. When searching for a job, workers exploit their social networks. We assume that information can flow within a network at zero cost and that there is no competition among workers belonging to the same network, as everyone can get the information at the same time. Thus, social networks can help workers find a job, and having a larger network represents an advantage for unemployed workers. This is reflected 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 with network $n^i_{o,t}$ can exert $(1 + n^i_{o,t})$ efficiency units of search.

$$U_{o,t} = \underbrace{\int \left(1 + n^i_{o,t}\right) \mathbb{1}\{e^i = 0; o^i = o\} di}_{\text{Search Effort of Unemployed workers}} + \underbrace{\int \xi \left(1 + n^{i,F}_{o,t}\right) \mathbb{1}\{e^i = 0; o^i = o^{i,F} = o; e^{i,F} = 1\} di}_{\text{Search Effort of Employed Fathers}} \tag{38}$$

where ξ captures possibly different levels of effectiveness of the father's network, compared to the personal network. The occupation–specific individual job finding rate $p^i_{o,t}$ is the sum of the probability of meeting a vacancy himself (either through his own effort or through his social network), and the probability of his father meeting a vacancy for him, if employed in the same occupation:

$$p^i_{o,t} = \frac{M_{o,t}}{U_{o,t}} \left[(1 + n^i_{o,t}) + \xi(1 + n^{i,F}_{o,t}) \mathbb{1}\left\{e^{i,F} = 1; o^{i,F} = o\right\}\right] \tag{39}$$

At each time t, workers face one single job–finding rate, in the occupation to which they are attached.

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 get exogenously destroyed at rate δ per period. When an employed worker gets separated from his match, he remains attached to his occupation.

Workers (both employed and unemployed) can freely reallocate across occupations. It is assumed that when an employed worker wants to reallocate, he is separated from the current match (the match is destroyed) and moves into the unemployment pool of his new occupation. Also, his occupation–specific human capital and social contacts stocks fully depreciate.

4.3 Intergenerational Transmission and Laws of Motion

We assume that upon entry in the labor market the duplet $\{\boldsymbol{\phi}^{\boldsymbol{P}}, \tau\}$ is imperfectly inherited from the father:

$$\boldsymbol{\phi}^{\boldsymbol{P}} \sim F\left(\boldsymbol{\phi}^{\boldsymbol{P}} | \boldsymbol{\phi}^{\boldsymbol{F,P}}\right) \tag{40}$$

$$\tau \sim G\left(\tau | \tau^{F}\right) \tag{41}$$

The initial level of occupation–specific human and social capital is assumed to be zero in all occupations. Both human and social capital, as well as preferences for occupations, evolve over time according to the following laws of motion:

$$h_{o,t+1} = F_h\left(h_{o,t} | e_t\right) \tag{42}$$

$$n_{o,t+1} = F_n\left(n_{o,t} | e_t\right) \tag{43}$$

Last, the temporary preferences vector is drawn each period from a distribution F_ϕ:

$$\boldsymbol{\phi}^{\boldsymbol{T}}{}_t \sim F_\phi \tag{44}$$

As already mentioned, upon changing occupation we assume that h_o and n_o start again from zero. We do this for computational reasons, even though in principle it would be interesting to keep track of all occupational–specific variables and make them decay over time if the worker is not attached to that occupation anymore.

4.4 Timing

The timing of our model is as follows:

1. Old (young) workers die (age) with probability ζ. The old ones who die are replaced by young unemployed workers.

2. Preference shocks are realized; reallocation of unemployed and employed workers takes place.

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

4. Exogenous separations take place. Unemployed workers find jobs or not.

5. The state variables of the workers are updated, according to the laws of motion.

4.5 Workers' Problem

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

Denote the state of a worker by $\Gamma = \{\boldsymbol{\phi}, \tau, h_o, n_o, o, e\}$, where o is set for simplicity equal to 0 in the case of a worker who has to choose an occupation for the first time. A young worker's choices are influenced by his father as he can get help in finding jobs from him, so that his own state includes also

his father's state Γ^F. We have to track all of father's state variables because offspring take into account that: i) even if unemployed (and therefore, not directly relevant for job–finding rate) today, their fathers will eventually help them find a job in the future; ii) fathers also change occupations over their life–cycle. In the following, we make explicit the dependance of workers' value functions on their employment status and occupation. Hence, conditional on employment status and occupational affiliation, we denote the state variable of young 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 down the value functions for old workers (denoted by a subscript F), with the understanding that old workers are characterized by $\Gamma^F = \varnothing$.

Fathers' problem

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

$$W_F^R(\Omega) = \max_{j \in \{1,\dots,O\}} \Big\{ W_{j,F}^U(\Omega) + \phi_{j,F}^T \Big\}, \tag{45}$$

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 deterministic component of the value of unemployment in occupation o, $W_{o,F}^U$, includes the value of unemployment benefits for the current period and the expected discounted value of the future[22]:

$$W_{o,F}^U(\Omega) = b + \tilde{\lambda} \Big[p_o(\Omega)\, \mathbb{E}\left[W_{o,F}^E(\Omega') \right] + (1 - p_o(\Omega))\, \mathbb{E}\left[W_{o,F}^R(\Omega') \right] \Big]. \tag{46}$$

An unemployed worker is matched with a vacancy in his occupation with probability $p_o(\Omega)$, and remains unemployed with the remainder $(1 - p_o(\Omega))$ probability, in which case he can decide to reallocate

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

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 reallocation decision at the beginning of the period. After that, if they decide to stay on the job, they receive the flow utility, earn the wage and keep with the same job next period, unless their match gets exogenously destroyed (this happens with probability δ). In this case, they will then choose their new occupation next period.

$$W^E_{o,F}(\Omega) = \max\Big\{\phi^P_{o,F} + \phi^{T,E}_{o,F} + w(\Omega) + \tilde{\lambda}\Big[(1-\delta)\,\mathbb{E}\left[W^E_{o,F}(\Omega')\right] + \delta\,\mathbb{E}\left[W^R_F(\Omega')\right]\Big], W^R_F(\Omega)\Big\}. \tag{47}$$

Note that employed workers draw a vector of size $(O+1)$–by–1 of preference shocks (that is, he draws a job–specific shock $\phi^{T,E}_{o,F}$ on top of the occupation–specific ones).[23]

4.6 Wage Determination

Upon matching, wages are set by a linear sharing rule, that is the worker is paid a fixed share of his productivity.[24] Let $y(\tau, h)$ be 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 are 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.

[23] This is done to preserve the symmetry between the occupational reallocation choices of unemployed and employed workers.

[24] Assuming other forms of wage determinations, such as Generalized Nash Bargaining, would be a rather extreme assumption in our setup, as 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.

4.7 Reallocation Across Occupations

We denote as j^* the preferred occupation for search, that is the occupation that maximizes the value of reallocation:

$$j_F^*(\Omega) \in \text{argmax}_{j\in\{1,\dots,O\}}\Big\{W_{j,F}^U(\Omega) + \phi_{j,F}^T\Big\}. \tag{48}$$

Next, let us define $\hat{W}_o^E(\Omega)$ the value of staying employed in occupation o (that is, the value of being employed and choosing not to reallocate):

$$\hat{W}_{o,F}^E(\Omega) = \phi_{o,F}^P + \phi_{o,F}^{T,E} + w(\Omega) + \tilde{\lambda}\Big[(1-\delta)\,\mathbb{E}\left[W_{o,F}^E(\Omega')\right] + \delta\,\mathbb{E}\left[W_F^R(\Omega')\right]\Big]. \tag{49}$$

Notice that $j^*(\Omega)$ might be the same or not as the current occupation of a worker. In case it is not, an unemployed worker will always want to reallocate, while for an employed worker the choice will depend on the difference between the value functions $\hat{W}_{o,F}^E(\Omega)$ and $W_F^R(\Omega)$.

We define $R_{o,F}^k(\Omega)$ (for $k \in \{E,U\}$) the policy function with respect to the reallocation decision: when $R_{o,F}^k(\Omega) = 1$, then a worker of type Ω with employment status k in occupation o optimally decides to reallocate. Optimality implies:

$$R_{o,F}^U(\Omega) = \mathbb{1}\{j_F^*(\Omega) \neq o\}.$$

$$R_{o,F}^E(\Omega) = \mathbb{1}\{W_F^R(\Omega) > \hat{W}_{o,F}^E(\Omega)\}.$$

Sons' problem

Sons face a very similar problem to the one of fathers. The only difference is that they do internalize their father's decisions. As a result, young workers can decide to reallocate as a consequence of a change in their own state variables (preferences) or because their father's state variables have changed, in which case they might want to follow their father in order to benefit from a higher probability of finding a job.

The expression for the value of reallocation stays exactly the same:

$$W_S^R(\Omega) = \max_{j \in \{1,\ldots,O\}} \left\{ W_{j,S}^U(\Omega) + \phi_{j,S}^T \right\}, \tag{50}$$

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

$$\begin{aligned} W_{o,S}^U(\Omega) = b + \lambda \Big[& p_o(\Omega) \left(\zeta \, \mathbb{E}\left[W_{o,F}^E(\Omega') \right] + (1-\zeta) \mathbb{E}\left[W_{o,S}^E(\Omega') \right] \right) \\ & + (1 - p_o(\Omega)) \left(\zeta \, \mathbb{E}\left[W_{o,F}^R(\Omega') \right] + (1-\zeta) \mathbb{E}\left[W_{o,S}^R(\Omega') \right] \right) \Big]. \end{aligned} \tag{51}$$

$$\begin{aligned} W_{o,S}^E(\Omega) = \max \Big\{ \phi_{o,S}^P + \phi_{o,S}^{T,E} + w(\Omega) + \lambda \Big[& (1-\delta) \left(\zeta \, \mathbb{E}\left[W_{o,F}^E(\Omega') \right] + (1-\zeta) \mathbb{E}\left[W_{o,S}^E(\Omega') \right] \right) \\ & + \delta \left(\zeta \, \mathbb{E}\left[W_F^R(\Omega') \right] + (1-\zeta) \mathbb{E}\left[W_S^R(\Omega') \right] \right) \Big], W_S^R(\Omega) \Big\}. \end{aligned} \tag{52}$$

Reallocation decisions $R_{o,S}^U(\Omega)$ and $R_{o,S}^E(\Omega)$ are isomorphic to the ones of the fathers, and are defined according to the above specified value functions.

4.8 Firms' Problem

Each firm is represented by one 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 holding the worker. The value

of holding an old worker reads as:

$$J_{o,F}(\Omega) = \mathbb{1}\{R^E_{o,F}(\Omega) = 0\}\Big[y(\tau, h_o) - w(\Omega) + \tilde{\lambda}\big[(1-\delta)\mathbb{E}\left[J_{o,F}(\Omega')\right] + \delta V'_o\big]\Big] + \mathbb{1}\{R^E_{o,F}(\Omega) = 1\}\, V_o. \tag{53}$$

The production of the match is given by the function $y(\tau, h_o)$, for which we assume y_1,$y_2 > 0$. Next period, with δ probability the match is exogenously destroyed (in which case, like in the case of the endogenous separation, the firm is left with the value of a vacancy V'_o). With the remainder $(1-\delta)$ probability, the match continues, and the state variables of the worker are updated.

Instead, the value of holding a young worker reads as:

$$J_{o,S}(\Omega) = \mathbb{1}\{R^E_{o,S}(\Omega) = 0\}\Big[y(\tau, h_o) - w(\Omega) + \lambda\big[(1-\delta)\left(\zeta\,\mathbb{E}\left[J_{o,F}(\Omega')\right] + (1-\zeta)\mathbb{E}\left[J_{o,S}(\Omega')\right]\right) + \delta V'_o\big]\Big] + \mathbb{1}\{R^E_{o,S}(\Omega) = 1\}\, V_o. \tag{54}$$

This equation has the same interpretation as the one for the old worker, except that it allows for the possibility that a worker turns old (when he is hit by the ζ shock) and the match continues.

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

$$V_o = -\kappa + q_o\,\mathbb{E}\left[J_o(\Omega')\right] + (1-q_o)V'_o, \tag{55}$$

where the expectation is taken over the distribution of unemployed workers attached to occupation o, both across all possible types Ω and possible ages $\{F, S\}$.

4.9 Worker Flows

The evolution of the stock of unemployed and employed workers is the result of optimal reallocation decisions, age shocks and labor market shocks (creation of new matches and exogenous destruction of existing ones). Let us define $g^k_{\Omega}(\tilde{\Omega}) = P(\Omega' = \Omega|\tilde{\Omega})$ the probability measure that a worker of type $\tilde{\Omega}$ with employment status k becomes of type Ω next 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.

Let us define $u'_{o,F}(\Omega)$ next period's measure of unemployed fathers of type Ω attached to occupation o:

$$u'_{o,F}(\Omega) = \int_{\Omega} \left[\hat{u}_{o,F}(\tilde{\Omega})\,(1 - R^U_{o,F}(\tilde{\Omega}))\,(1 - p_{o,F}(\tilde{\Omega}))\,g^U_{\Omega}(\tilde{\Omega}) + \hat{e}_{o,F}(\tilde{\Omega})\,(1 - R^E_{o,F}(\tilde{\Omega}))\,\delta\,g^E_{\Omega}(\tilde{\Omega})\right] d\tilde{\Omega} \quad (56)$$
$$+ \sum_{\tilde{o}\neq o} \int_{\Omega} \left[\hat{u}_{\tilde{o},F}(\tilde{\Omega})\,R^U_{\tilde{o},F}(\tilde{\Omega}) + \hat{e}_{\tilde{o},F}(\tilde{\Omega})\,R^E_{\tilde{o},F}(\tilde{\Omega})\right] \mathbb{1}\{j^*_F(\tilde{\Omega}) = o\}\,(1 - p_{o,F}(\tilde{\Omega}))\,g^U_{\Omega}(\tilde{\Omega})\,d\tilde{\Omega},$$

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 this stock includes fathers who do not die, as well as sons who become fathers.

Equation 56 is composed of four different terms: the first two refer respectively to unemployed workers attached to occupation o who decided not to reallocate and did not find a job, and employed workers in occupation o who did not reallocate and lost their job; the last two terms are (unemployed and employed) workers who decided to reallocate into occupation o but did not find a job in the previous period.

Turning to employed fathers, $e'_o(\Omega)$ is defined as follows:

$$e'_{o,F}(\Omega) = \int_{\Omega} \left[\hat{e}_{o,F}(\tilde{\Omega}) \left(1 - R^E_{o,F}(\tilde{\Omega})\right) (1-\delta)\, g^E_{\Omega}(\tilde{\Omega}) + \hat{u}_{o,F}(\tilde{\Omega}) \left(1 - R^U_{o,F}(\tilde{\Omega})\right) p_{o,F}(\tilde{\Omega})\, g^U_{\Omega}(\tilde{\Omega}) \right] d\tilde{\Omega} \quad (57)$$
$$+ \sum_{\tilde{o} \neq o} \int_{\Omega} \left[\hat{u}_{\tilde{o},F}(\tilde{\Omega})\, R^U_{\tilde{o},F}(\tilde{\Omega}) + \hat{e}_{\tilde{o},F}(\tilde{\Omega})\, R^E_{\tilde{o},F}(\tilde{\Omega}) \right] \mathbb{1}\{j^*_F(\tilde{\Omega}) = o\}\, p_{o,F}(\tilde{\Omega})\, g^U_{\Omega}(\tilde{\Omega})\, d\tilde{\Omega}.$$

The stock of employed is given by workers who were already employed last period in the same occupation and neither did not lose the job nor found profitable to reallocate, plus the mass of unemployed workers who did not want to reallocate and found a vacancy, plus all workers who had just reallocated into occupation o and found a job.

For what concerns the distribution of sons, the one of the employed is exactly symmetric to the fathers':

$$e'_{o,S}(\Omega) = \int_{\Omega} \left[(1-\zeta) e_{o,S}(\tilde{\Omega}) \left(1 - R^E_{o,S}(\tilde{\Omega})\right) (1-\delta)\, g^E_{\Omega}(\tilde{\Omega}) + (1-\zeta) u_{o,S}(\tilde{\Omega}) \left(1 - R^U_{o,S}(\tilde{\Omega})\right) p_{o,S}(\tilde{\Omega})\, g^U_{\Omega}(\tilde{\Omega}) \right] d\tilde{\Omega} \quad (58)$$
$$+ \sum_{\tilde{o} \neq o} \int_{\Omega} \left[(1-\zeta) u_{\tilde{o},S}(\tilde{\Omega})\, R^U_{\tilde{o},S}(\tilde{\Omega}) + (1-\zeta) e_{\tilde{o},S}(\tilde{\Omega})\, R^E_{\tilde{o},S}(\tilde{\Omega}) \right] \mathbb{1}\{j^*_S(\tilde{\Omega}) = o\}\, p_{o,S}(\tilde{\Omega})\, g^U_{\Omega}(\tilde{\Omega})\, d\tilde{\Omega}.$$

Finally, the distribution of unemployed sons reads as follows:

$$u'_{o,S}(\Omega) = \int_{\Omega} \left[(1-\zeta) u_{o,S}(\tilde{\Omega}) \left(1 - R^U_{o,S}(\tilde{\Omega})\right) \left(1 - p_{o,S}(\tilde{\Omega})\right) g^U_{\Omega}(\tilde{\Omega}) + (1-\zeta) e_{o,S}(\tilde{\Omega}) \left(1 - R^E_{o,S}(\tilde{\Omega})\right) \delta\, g^E_{\Omega}(\tilde{\Omega}) \right] d\tilde{\Omega} \quad (59)$$
$$+ \sum_{\tilde{o} \neq o} \int_{\Omega} \left[(1-\zeta) u_{\tilde{o},S}(\tilde{\Omega})\, R^U_{\tilde{o},S}(\tilde{\Omega}) + (1-\zeta) e_{\tilde{o},S}(\tilde{\Omega})\, R^E_{\tilde{o},S}(\tilde{\Omega}) \right] \mathbb{1}\{j^*_S(\tilde{\Omega}) = o\} \left(1 - p_{o,S}(\tilde{\Omega})\right) g^U_{\Omega}(\tilde{\Omega})\, d\tilde{\Omega}$$
$$+ \zeta \frac{\mathbb{1}\{\Omega \in \Omega^{NB}\}}{\int_{\Omega} \mathbb{1}\{\Omega \in \Omega^{NB}\}},$$

where the only difference comes from the last term, which represent the flow of newborns, randomly

directed to the subset Ω^{NB} of the entire state space.

4.10 Equilibrium Definition

We focus on a symmetric Steady State equilibrium (SS) in which all value functions and reallocation decisions of workers are constant over time. In particular, value functions and optimal decisions only depend on the individual state Ω.

Definition: A 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, reallocation 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:

- Value functions for workers follow from equations 46, 47, 51, 52 and reallocation decisions are optimal for them.
- The value of a vacancy is equal to zero in all occupations: $V_o = 0$ (that is, there is free–entry in all occupations).
- Labor market tightness is consistent with free–entry and pinned down by equation 55.
- Wages are pinned down according to equation 35.
- Individual state variables evolve according to equations 42, 43 and 44.
- Distributions of workers evolve according to equations 56, 57, 58 and 59.
- 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 our model (ability, preferences and networks transmission) in delivering occupational persistence. First, we need to assign values to the structural parameters of our model. Next, we will use the calibrated model to decompose occupational persistence and to perform welfare analysis and policy experiments.

5.1 Identification Strategy

Our strategy involves fixing exogenously some of the parameters, and jointly calibrating all the rest to relevant moments of the UK data. First of all, we fix the number of occupations O to 4.[25] One period in the model corresponds to one month, hence the discount factor λ is fixed to 0.9966. The age shock ζ is set to deliver an average worklife of 40 years (20 as young and 20 as old), implying a value of 0.00416. Last, we fix the surplus sharing rule parameter χ to 0.7, and the parameters governing the matching function A and η to 0.1 and 0.5 (as in Petrongolo and Pissarides 2001), respectively.

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 lowest 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 come from tenure. The accumulation/depreciation of these occupation–specific variables are subject to a Markov-process characterized by the following parameters: p_h^+, p_h^-, p_n^+, p_n^-, where the $+$ and $-$ superscripts denote accumulation

[25] The calibration with 9 occupations (resembling the 1–digit level aggregation) is work in progress.

(when employed) and depreciation (when unemployed) probabilities, respectively.

We also assume that each worker has a talent τ in one occupation, where he gets a productivity premium $\hat{\tau}$. The minimum level of productivity is normalized to 1. The match 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 $\bar{\phi} = 0$. Both τ and ϕ are drawn at birth, with ρ_τ and ρ_ϕ being the probability of drawing the same as the father (values of ρ's equal to 1 represent perfect persistence). Let τ and ϕ be the indicators of the occupation in which a worker has the comparative advantage and the preference premium, respectively:

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

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

Finally, we assume that the idiosyncratic preference shocks are drawn from an Extreme distribution of Type-1, with standard deviation σ.

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 pick $K = 14$, that is the model is exactly identified). Table 7 contains the list of all parameters of the model, each of which is associated to 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 means that sons can take advantage of a larger share of their father's network. The transmission of comparative advantage is calibrated to match the intergenerational wage correlation (from Corak (2006)), that is equal to 0.5 in the UK. An increase in ρ_τ determines an increase in the wage correlation of followers and a decrease in the wage correlation of movers, as well as an increase in the share of followers. Instead, 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 match entirely the 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 idea behind this is that the more heterogeneous the productivity levels of workers across occupations are, the more dispersed equilibrium wages will be. The networks premium $\hat{n}$ is calibrated to match the share of jobs found through networks in the UK, which is 0.23 (Pellizzari

(2010)). The higher $\hat{n}$, the more networks will be present in the economy and used for job search. The preference premium $\hat{\phi}$ is chosen to make the model 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 for the occupational choice. This is necessary, in order for the model to be able to induce some workers to become followers for reasons different from productivity. 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 frequent occupational changes are. 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 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 month 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 match all targets precisely, both the cross–sectional and the inter–generational ones. We can replicate the entire amount of occupational persistence observed in the data, by making preferences quite persistent across generations (0.48), substantially more than the comparative advantages are (0.36). The share of parental networks exploited by offspring is 0.37, which generates the same job–finding rate premium as in the data.

Moreover, a large degree of heterogeneity is needed in order to match the data moments. The

preference premium is 0.82, while the comparative advantage premium is even higher, 1.03. The networks premium is also very substantial (0.93), whereas the human capital premium is 0.26 (directly taken from the data). The accumulation of both human capital and networks is slow, while their depreciation over unemployment is very fast, especially for human capital.

Search frictions are substantial, as posting a vacancy costs 5.37 units of the consumption good. 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	0.369	JF premium of followers (BHPS)	0.055	0.055
ρ_τ	Transmission of Comparative Advantage	0.360	Interg. Wage correlation (Corak 2006)	0.500	0.498
ρ_ϕ	Transmission of Preferences	0.483	Interg. Occupational Persistence (Lik. Ratio, BHPS)	1.720	1.719
Heterogeneity and Laws of Motions					
$\hat{h}$	HC premium	0.260	Average occupational tenure returns after 5 years (BHPS)	-	-
$\hat{n}$	Networks premium	0.932	Share of jobs found through contacts (Pellizzari 2010)	0.230	0.230
$\hat{\tau}$	Comparative Advantage premium	1.029	Within–occupation log wage variance (BHPS)	0.166	0.166
$\bar{\phi}$	Baseline preference for jobs	0	Normalization	-	-
$\hat{\phi}$	Preference premium	0.820	Wage discount of followers (BHPS)	0.076	0.076
p_h^+	Prob. of accumulating HC (employed)	0.0166	Avg. occupational tenure returns after 5 years (BHPS)	0.260	0.260
p_h^-	Prob. of losing HC (unemployed)	0.998	Avg. wage discount after unemp. (Arulampalam 2001)	0.076	0.076
p_n^+	Prob. of accumulating networks (employed)	0.006	Regression of JF rate vs. past occupational tenure (BHPS)	0.008	0.008
p_n^-	Prob. of losing networks (unemployed)	0.206	JF rate–unemployment duration profile (see text, BHPS)	1.066	1.075
σ	Standard deviation of preference shocks	0.423	Occupational change rate, after U spell (monthly, BHPS)	0.357	0.355
Environment					
O	Number of occupations	4	Normalization	-	-
κ	Vacancy posting cost	5.366	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	0.756	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	Normalization	-	-

Table 7. Calibration Results

5.3 Occupational Persistence Decomposition and Welfare Analysis

The model allows us to study what are the factors behind the occupational choice, and how this possibly varies between followers and movers. In Table 8, we calculate how often the occupational choice is aligned with each of the three possible factors (parental networks, comparative advantage, preferences) under the baseline calibration.

	All	**Followers**	**Movers**
Sorting along comparative advantage (fathers)	0.656	-	-
Sorting along preferences (fathers)	0.594	-	-
Sorting along parental networks (sons)	0.430	1.000	0.000
Sorting along comparative advantage (sons)	0.726	0.668	0.770
Sorting along preferences (sons)	0.524	0.573	0.486
Average log wage (sons)	0.311	0.268	0.344
Average unemployment rate (sons)	0.061	0.048	0.071

Table 8. Occupational sorting.

For fathers, comparative advantages seem to be more important than preferences for the occupational sorting (respectively, about 66% and 59% of them choose the occupation aligned with that specific factor). Among sons, the same holds true: about 73% of them pick the occupation in which they are most productive, whereas about 52% of them pick their preferred occupation; finally, the occupational choice is aligned with parental networks for all followers (by construction), that is in 43% of the cases. Striking differences in the sorting arises between followers and movers: the former put more weight on preferences in their occupational decision (57% against 49% of movers) and less on their productivity (67% as opposed to 77% of movers). As a consequence of this, followers earn lower wages, as shown in row 4 of Table 8. At the same time, followers have better employment prospects than movers, with an average unemployment rate of 4.8%, against 7.1% of movers. Summing up, the model economy generates a clear sorting of workers in the two regions high–employment/low–wages, low–employment/high–wages.

However suggestive, these correlations are not yet informative on the nature of occupational persistence. For this reason, we now perform the exercise of sequentially shutting down each of the 3 channels delivering occupational persistence. In this way, we are able to: i) quantify the contributions of each channel to the overall persistence; ii) evaluate welfare in the different scenarios. Table 9 shows the results of the different experiments. In column 1 we have our 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. Shutting down parental networks generates a drop in persistence of about 59% (column 2), while comparative advantages and preferences transmission on their own account for about 26% and 23% (columns 3 and 4) of persistence, respectively. Moreover, networks and preferences transmission seem to reinforce each other as a source of persistence, while comparative advantages and preferences transmission are more independent forces, as shutting down these channels in pairs delivers less (more) of a drop than the sum of the effects separately (columns 5–7 vs 2–4).

Second, the welfare consequences of a reduction in persistence are very different across the different experiments. When we shut down parental networks (column 2), welfare improves by 0.1%, due to the better allocation of workers to occupations (the sorting along productivity increases from 73% to 77%) and despite a worse sorting along preferences (it drops from 52% to 48%). As a consequence of the increase in the productivity of the workforce, output per worker goes up and wage variance goes down. Also, despite the fact that less efficiency units of search are now exerted in the market, unemployment improves, as firms react to the change in average labor productivity by posting more vacancies. On the contrary, when we shut down the transmission of comparative advantages (column 2), welfare decreases by 0.2%, with declining output per worker (sorting along productivity decreases, the one along preferences improves) and rising unemployment (by 1.3%). Last, shutting down preferences transmission (column 3) has a very similar effect to shutting down parental networks: productivity becomes more prominent in individuals' choice, output per worker increases and unemployment decreases. The net

effect of these changes, despite a worse sorting along preferences, is an increase in welfare of 0.2%.

		(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	1.294	1.535	1.552	1.104	1.190	1.349	1.000
(rel. change from baseline)		-	(-0.590)	(-0.255)	(-0.232)	(-0.856)	(-0.736)	(-0.514)	(-1.000)
Welfare (rel. change from baseline)		-	(0.001)	(-0.002)	(0.002)	(0.001)	(0.001)	(-0.000)	(0.001)
Sorting along comparative advantage (sons)		0.726	0.772	0.709	0.754	0.772	0.772	0.738	0.772
Sorting along preferences (sons)		0.524	0.477	0.541	0.495	0.477	0.477	0.511	0.478
Sorting along comparative advantage (fathers)		0.656	0.656	0.656	0.656	0.656	0.656	0.656	0.656
Sorting along preferences (fathers)		0.594	0.594	0.594	0.594	0.594	0.594	0.594	0.594
Output per worker (=1 in baseline)		1.000	1.027	0.990	1.017	1.027	1.027	1.007	1.027
Log Wage Variance (rel. change from baseline)		0.166	(-0.065)	(0.021)	(-0.038)	(-0.065)	(-0.065)	(-0.016)	(-0.065)
Welfare CV (rel. change from baseline)		0.190	(-0.002)	(0.001)	(-0.002)	(-0.002)	(-0.002)	(-0.001)	(-0.002)
Unemployment Rate (rel. change from baseline)		0.061	(-0.018)	(0.013)	(-0.013)	(-0.018)	(-0.018)	(-0.000)	(-0.018)
Average UE Rate (rel. change from baseline)		0.125	(0.025)	(-0.011)	(0.015)	(0.025)	(0.025)	(0.004)	(0.025)
Average EU Rate (rel. change from baseline)		0.005	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.008)	(0.003)
Equilibrium Tightness		1.564	(0.051)	(-0.022)	(0.030)	(0.051)	(0.051)	(0.008)	(0.051)

Table 9. Occupational Persistence Decomposition and Welfare Analysis.

5.4 Other Counterfactual Experiments

5.4.1 The Importance of Multiple Transmission Channels

Our model is characterized by three levels of heterogeneity, and intergenerational persistence concerns all levels of heterogeneity considered here. One interesting exercise 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 for replicating the data patterns which are the targets of our baseline calibration. We focus on the transmission of productive abilities as the only transmission channel, because this channel embeds in a reduced form way genetic transmission, educational choices and HC transmission 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 7 except for the JF rate premium and the intergenerational wage correlation.

In general, the fit of the model has substantially worsened, as we can see from Table 10. The model is not able to fully account for occupational persistence, producing a likelihood ratio of 1.62. The value of ρ_τ is higher than in the baseline calibration, as this is now the only source of occupational persistence. As a consequence of this, the model largely undershoots log wage variance, which is only 0.022. Importantly, the model completely fails at generating the wage discount between followers and movers. This reflects the fact that with only one source of persistence it is impossible to induce different sorting behaviors among agents. By construction, the model also cannot replicate the job–finding rate premium of followers (non–targeted), as any networks transmission is shut down. Finally, the intergenerational wage correlation (non–targeted) implied in the model is 0.35, substantially lower than the one observed in the data (0.50).

When we perform the exercise of shutting down occupational persistence in the restricted economy (table 11), 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,

that vanishes completely. This is because in this economy persistence is not a sign of any distortion in the occupational choice of individuals. Besides, persistence in this economy is generated only by the fact that father–son couples tend to be more similar than two randomly picked workers. In this sense, occupational persistence is not a reflection of the fact that offspring *care* about the occupational choices of their fathers and are affected by them. In other words, offspring's policy and value functions are now independent of the father's state variables.

Parameter	Description	Value	Target/Source	Data	Model
Intergenerational Transmission					
ξ	Transmission of Networks	0*	JF premium of followers (BHPS)	0.055	0.000
ρ_τ	Transmission of Comparative Advantage	0.429	Interg. Wage correlation (Corak 2006)	0.500	0.347
ρ_ϕ	Transmission of Preferences	0.250*	Interg. Occupational Persistence (Lik. Ratio, BHPS)	1.720	1.616
Heterogeneity and Laws of Motions					
$\hat{h}$	HC premium	0.260	Average occupational tenure returns after 5 years (BHPS)	-	-
$\hat{n}$	Networks premium	1.721	Share of jobs found through contacts (Pellizzari 2010)	0.230	0.213
$\hat{\tau}$	Comparative Advantage premium	0.331	Within–occupation log wage variance (BHPS)	0.166	0.022
$\bar{\phi}$	Baseline preference for jobs	0	Normalization	-	-
$\hat{\phi}$	Preference premium	0.230	Wage discount of followers (BHPS)	0.076	0.001
p_h^+	Prob. of accumulating HC (employed)	0.0166	Avg. occupational tenure returns after 5 years (BHPS)	0.260	0.260
p_h^-	Prob. of losing HC (unemployed)	0.828	Avg. wage discount after unemp. (Arulampalam 2001)	0.076	0.068
p_n^+	Prob. of accumulating networks (employed)	0.003	Regression of JF rate vs. past occupational tenure (BHPS)	0.008	0.008
p_n^-	Prob. of losing networks (unemployed)	0.146	JF rate–unemployment duration profile (see text, BHPS)	1.066	1.088
σ	Standard deviation of preference shocks	0.220	Occupational change rate, after U spell (monthly, BHPS)	0.357	0.307
Environment					
O	Number of occupations	4	Normalization	-	-
κ	Vacancy posting cost	5.520	Average UE rate (monthly, BHPS)	0.125	0.092
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	0.553	Average replacement rate (OECD)	0.530	0.511
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 10. Calibration Results, Restricted Model (* = restricted parameters; [] =non–targeted moments).

		(1)	(2)
No parental net.	$(\xi = 0)$	n.a.	n.a.
No comp. adv. trans.	$(\rho_\tau = 1/O)$	-	✓
No pref. trans.	$(\rho_\phi = 1/O)$	n.a.	n.a.
Occupational Persistence		1.616	1.000
(Δ% from baseline)		-	(-1.000)
Welfare (rel. change from baseline)		-	(-0.000)
Sorting along comparative advantage (sons)		0.997	0.997
Sorting along preferences (sons)		0.253	0.253
Sorting along comparative advantage (fathers)		0.781	0.781
Sorting along preferences (fathers)		0.469	0.469
Output per worker (=1 in baseline)		1.000	1.000
Log Wage Variance (rel. change from baseline)		0.022	(0.000)
Welfare CV (rel. changefrom baseline)		0.091	(0.000)
Unemployment Rate (rel. change from baseline)		0.083	(0.000)
Average UE Rate (rel. change from baseline)		0.092	(-0.000)
Average EU Rate (rel. change from baseline)		0.005	(0.000)
Equilibrium Tightness		0.850	(-0.000)

Table 11. Occupational Persistence Decomposition and Welfare Analysis (Restricted Model).

5.4.2 The Role of Search Frictions

We investigate the role of search frictions, imposing on the UK baseline calibration the degree of frictions implied by the monthly finding rates of different economies. We focus on two polar cases among OECD countries, US and Spain. We take the average finding rates estimated in Hobijn & Şahin (2009), who report 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 0.83 and 15.61, respectively. We repeat the exercises of column 2–4 of the previous table for both of the counterfactual economies, with the results shown in Table 12.

Two main results stand out. First, the importance of parental networks crucially depends on the size of the frictions. In the low–frictions economy, removing the effect of parental networks barely affects persistence (it drops by only 1.4%), whereas the drop in the high–frictions economy is much more pronounced (57%). At the same time, the removal of networks is welfare improving in the high–frictions economy (as it raises average labor productivity), but it is welfare decreasing in the low–frictions economy (it crowds out occupational choice along preferences, as networks are basically not generating any occupational choice not based on productivity in the baseline equilibrium).

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

We repeat the same counterfactual exercise under the restricted calibration (in columns 9 to 12). The comparison between column 9 and column 11 reveals that the same conclusions hold true: high levels of frictions cause a worse occupational sorting, and as a consequence a lower level of output per worker and a higher level of unemployment. Even though the transmission of contacts is shut down, changes in the level of frictions interact with how easily workers are willing to reallocate across occupations. That is, when a worker finds himself in an occupation in which he does not have the comparative advantage, he would find it very unattractive to change occupation (a decision which entails an unemployment spell) in order to pursue high wages. As a consequence, sorting along productivity worsens in the high–frictions economy. Notice that the productivity differential between the two economies is 6.5%, substantially smaller than the differential between the two economies under the baseline calibration (about 21%). The absence of intergenerational transmission of networks of the restricted calibration economy implies that search frictions can be harmful to the allocation of the labor force to a much lesser extent than in the baseline economy.

		Baseline Economy ($\kappa = 0.83$)				Baseline Economy ($\kappa = 15.61$)				Restricted Model ($\kappa = 0.83$)		Restricted Model ($\kappa = 15.61$)	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
No parental net.	($\xi = 0$)	-	✓	-	-	-	✓	-	-	n.a.	n.a.	n.a.	n.a.
No comp. adv. trans.	($\rho_\tau = 1/O$)	-	-	✓	-	-	-	✓	-	-	✓	-	✓
No pref. trans.	($\rho_\phi = 1/O$)	-	-	-	✓	-	-	-	✓	n.a.	n.a.	n.a.	n.a.
Occupational Persistence		1.436	1.429	1.006	1.434	1.719	1.310	1.584	1.488	1.711	1.000	1.285	1.000
		-	(-0.014)	(-0.987)	(-0.002)	-	(-0.568)	(-0.187)	(-0.321)	-	(-1.000)	-	(-1.000)
Welfare		-	(-0.008)	(-0.001)	(0.000)	-	(0.002)	(-0.002)	(0.002)	-	(-0.000)	-	(-0.000)
Sorting along comp. adv. (sons)		0.993	0.993	0.993	0.993	0.648	0.672	0.633	0.677	0.998	0.998	0.758	0.758
Sorting along pref. (sons)		0.256	0.256	0.256	0.256	0.602	0.579	0.617	0.573	0.251	0.251	0.491	0.491
Sorting along comp. adv. (fathers)		0.993	0.993	0.993	0.993	0.635	0.635	0.635	0.635	0.999	0.999	0.681	0.681
Sorting along pref. (fathers)		0.257	0.257	0.257	0.257	0.615	0.615	0.615	0.615	0.251	0.251	0.567	0.567
Output per worker		2.371	(0.000)	(-0.000)	(-0.000)	1.956	(0.015)	(-0.009)	(0.018)	1.570	(-0.000)	1.473	(-0.000)
Log Wage Variance		0.022	(-0.006)	(0.009)	(-0.000)	0.167	(-0.018)	(0.009)	(-0.022)	0.015	(0.000)	0.034	(0.000)
Welfare CV		0.190	(-0.001)	(-0.000)	(0.000)	0.192	(-0.000)	(-0.000)	(-0.003)	0.087	(0.000)	0.107	(0.000)
Unemployment Rate		0.014	(0.002)	(-0.019)	(0.000)	0.182	(-0.010)	(0.011)	(-0.011)	0.014	(0.000)	0.212	(0.000)
Average UE Rate		0.565	(0.019)	(0.013)	(-0.000)	0.040	(0.015)	(-0.011)	(0.016)	0.460	(-0.000)	0.031	(-0.000)
Average EU Rate		0.012	(-0.081)	(0.006)	(0.002)	0.005	(0.000)	(0.004)	(0.004)	0.007	(-0.000)	0.004	(0.000)
Equilibrium Tightness		32.509	(0.034)	(0.024)	(-0.001)	0.158	(0.029)	(-0.022)	(0.031)	22.041	(-0.000)	0.098	(-0.000)

Table 12. 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 labor market frictions; Columns (5)-(8) are under the baseline calibration and Spain labor market frictions; Columns (9)-(10) are under the calibration of the restricted model and the US labor market frictions; Columns (5)-(8) are under the calibration of the restricted model and Spain labor market frictions.

5.4.3 Policy Experiment: Unemployment Benefits

We now study how changes in unemployment benefits affect the equilibrium of our model economy. In order to assess the welfare consequences of such changes, we introduce lump–sum taxation τ on existing matches (split among workers and firms, where χ is the share paid by the workers) and the government BC. The new value functions for employed workers and firms read as follows:

$$W_o^E(\Omega) = \max\Big\{\phi_o^P + \phi_o^{T,E} + w(\Omega) - \chi\tau + \tilde{\lambda}\Big[(1-\delta)\,\mathbb{E}\left[W_o^E(\Omega')\right] + \delta\,\mathbb{E}\left[W^R(\Omega')\right]\Big], W^R(\Omega)\Big\}. \quad (60)$$

$$J_o(\Omega) = \mathbb{1}\{R_o^E(\Omega) = 0\}\Big[y(\tau, h_o) - w(\Omega) - (1-\chi)\tau + \tilde{\lambda}\big[(1-\delta)\mathbb{E}\left[J_o(\Omega')\right] + \delta V_o'\big]\Big] + \quad (61)$$
$$\mathbb{1}\{R_o^E(\Omega) = 1\}\, V_o.$$

The government is balancing the budget at each period. That is, the change in unemployment benefits from the baseline equilibrium needs to be financed with the tax revenues:

$$\Delta b\, u = \tau(1-u), \quad (62)$$

where u is the unemployment rate of the economy. The rest of the model stays the same.

Some of the channels commonly emphasized by 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 trade–off at work in our model. An increase (decrease) in the value of unemployment benefits decreases (increases) the distance between the value of employment and unemployment. As a consequence, workers are more (less) prone to accept an unemployment spell, and therefore will be better (worse) matched in terms of productivity, on average.

Our mechanism has a very similar flavor to the one in Acemoglu & Shimer (2000) and Golosov *et al.* (2013).

As we can see from the upper part of Table 13, an increase of 10% (25%) in b causes output per worker to increase by about 0.1% (0.4%), whereas decreasing b causes output per worker to fall. At the same time, unemployment goes up when we increase b (we have both more endogenous separations and smaller finding rates), and goes down otherwise. This directly follows from the fact that unemployment is now a more (less) attractive option.

The overall net effect on welfare is negative in the case of increases in b, and positive in the case of decreases. This reflects the fact that the tax rate increases more than proportionally with increases in b, given that unemployment goes up. Notice that the change in τ is such that, in face of the improvement of the quality of the labor force, firms' entry goes down. When we repeat the same exercise under the restricted calibration (lower part of Table 13), we find that effects are much larger. This is because the sorting along productivity is already almost perfect, and increases in b bring about negligible improvements in output per worker. At the same time, the temporary preference shocks become more prevalent in occupational choice, contributing to a large spike in unemployment (it goes up by more than 29%, when we increase b by 25%). As a consequence, the small improvements in output per worker mitigate the drop in welfare due to increased taxation to a much lesser extent than before. For instance, the welfare cost of increasing b by 25% is 0.9%, as opposed to 0.1% in the baseline economy. Hence, it turns out that having multiple sources of persistence is relevant for the assessment of labor market policy in general.

6 Conclusions

We have investigated the determinants of occupational persistence across generations. When persistence comes from multiple sources, it is key to assess their relative importance, in order to understand the

Change in b (Δ% from baseline)	(1) -	(2) +10%	(3) +25%	(4) -10%	(5) -25%
Baseline Economy					
Occupational Persistence	1.719	1.702	1.675	1.735	1.758
(rel. change from baseline)	-	(-0.024)	(-0.061)	(0.022)	(0.054)
Welfare (rel. change from baseline)	-	(-0.000)	(-0.001)	(0.000)	(0.000)
Sorting along comparative advantage (sons)	0.726	0.728	0.733	0.724	0.721
Sorting along preferences (sons)	0.524	0.521	0.516	0.526	0.529
Sorting along comparative advantage (fathers)	0.656	0.656	0.656	0.656	0.656
Sorting along preferences (fathers)	0.594	0.594	0.594	0.594	0.594
Output per worker (=1 in baseline)	1.000	1.001	1.004	0.999	0.997
Log Wage Variance (rel. change from baseline)	0.166	(0.002)	(0.002)	(-0.002)	(-0.005)
Welfare CV (rel. change from baseline)	0.190	(-0.000)	(-0.001)	(0.000)	(0.001)
Unemployment Rate (rel. change from baseline)	0.061	(0.026)	(0.075)	(-0.022)	(-0.048)
Average UE Rate (rel. change from baseline)	0.125	(-0.013)	(-0.037)	(0.012)	(0.026)
Average EU Rate (rel. change from baseline)	0.005	(0.025)	(0.072)	(-0.021)	(-0.046)
Equilibrium Tightness	1.564	(-0.027)	(-0.073)	(0.023)	(0.054)
Restricted Model					
Occupational Persistence	1.616	1.611	1.599	1.543	1.549
(rel. change from baseline)	-	(-0.008)	(-0.026)	(-0.119)	(-0.108)
Welfare (rel. change from baseline)	-	(-0.002)	(-0.009)	(0.001)	(0.002)
Sorting along comparative advantage (sons)	0.997	0.997	0.995	0.939	0.942
Sorting along preferences (sons)	0.253	0.254	0.256	0.311	0.308
Sorting along comparative advantage (fathers)	0.781	0.778	0.777	0.783	0.786
Sorting along preferences (fathers)	0.469	0.472	0.472	0.466	0.464
Output per worker (=1 in baseline)	1.000	1.000	0.999	0.986	0.986
Log Wage Variance (rel. change from baseline)	0.022	(0.013)	(0.034)	(0.176)	(0.153)
Welfare CV (rel. change from baseline)	0.091	(-0.001)	(-0.002)	(0.008)	(0.011)
Unemployment Rate (rel. change from baseline)	0.083	(0.085)	(0.293)	(-0.107)	(-0.149)
Average UE Rate (rel. change from baseline)	0.092	(-0.045)	(-0.135)	(0.056)	(0.086)
Average EU Rate (rel. change from baseline)	0.005	(0.082)	(0.272)	(-0.123)	(-0.163)
Equilibrium Tightness	0.850	(-0.087)	(-0.251)	(0.115)	(0.179)

Table 13. Policy Experiment: effect of changes in unemployment benefits.

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 have extended our 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 have found that all channels are quantitatively relevant, and that a model based only on transmission of ability (restricted model) would be at odds with several dimensions of the data.

We have also studied how changes in search frictions can cause at the same time large differentials in both output per worker and unemployment rates. Importantly, the restricted model would largely understate the harmful effects of frictions on the allocation of workers to occupations. When we study the cost of increasing unemployment benefits, the restricted model overstates such costs by almost an order of magnitude. Hence, we conclude that modeling multiple sources of intergenerational transmission is crucial not only to understand the consequences of persistence, but also for the assessment of labor market policy in general.

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**.

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.*

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.

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

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.

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.

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 womenâs 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. 2013. 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.

Papageorgiou, Theodore. 2013. Learning your comparative advantages. *The Review of Economic Studies*, rdt048.

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*.

7 Appendix

Proofs

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 to: $-\Big[(1-\Psi)[(1+N)U^* - \mu(1+N)\frac{\partial U^*}{\partial \mu}] - \mu(1+N)\frac{\partial \Psi^*}{\partial \mu}U^*\Big]$. The second term is clearly non–positive, as $\frac{\partial \Psi^*}{\partial \mu} \leq 0$. For what concerns the first term, it suffices to show that: $U^* \geq \mu\frac{\partial U^*}{\partial \mu}$. The left–hand side reads as follows: $1+\mu N+\Psi(1-\mu)n$, 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, as $\frac{\partial \Psi^*}{\partial \mu} \leq 0$. 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:

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

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

Evaluating the free–entry condition at equilibrium:

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

Given that $q'(\theta) < 0$, this 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

Evaluating persistence at equilibrium:

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

Now take 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 one is always positive. We can show that even in the case $\mu = 0$ (that is, when we make the negative term as large as possible), the sum of the two is still positive. To see this, using 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, $\frac{\partial \mathcal{P}^*}{\partial \mu} \geq 0$. 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 $\frac{\partial \Psi^*}{\partial \rho} \geq 0$.

Proposition 2

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). Here we provide a sufficient condition for the direct effect of an increase in μ (ρ) to be also negative (positive). Look at the direct effect of μ:

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

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

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

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

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

One can easily verify 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, $\frac{d\mathcal{W}}{d\mu} \leq 0$.

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

Additional Empirical Evidence

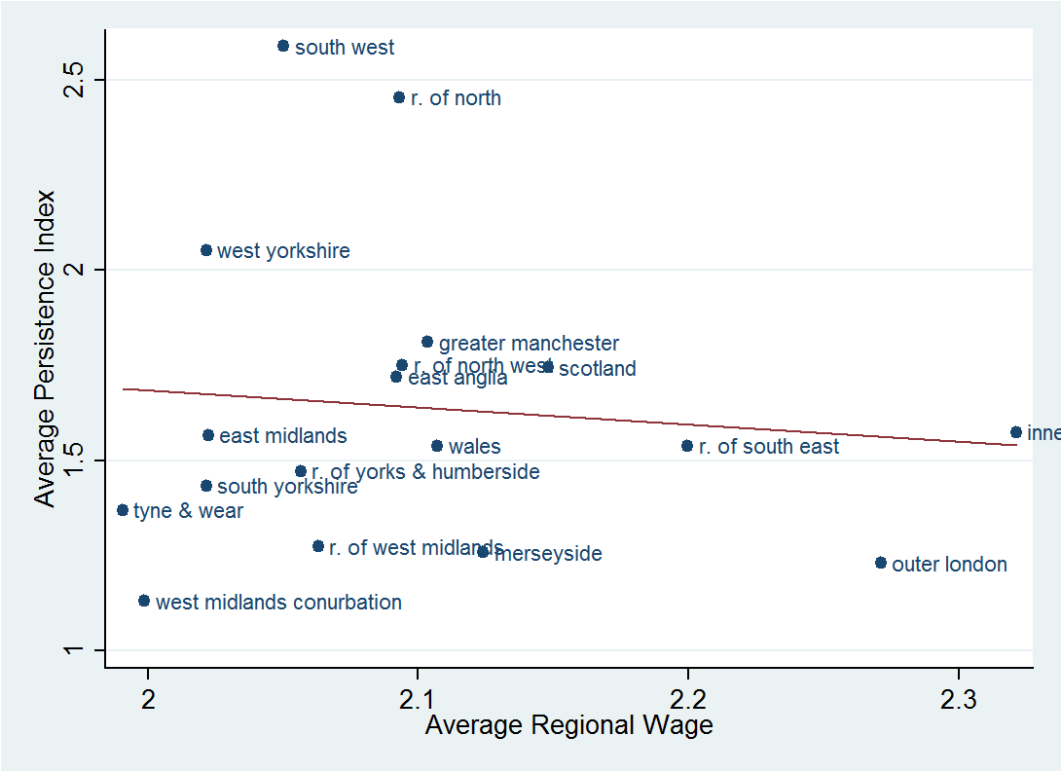

Figure 4. Plot of Average Persistence Index (weighted average of occupation–specific likelihood ratios, by region) vs. Average Regional Wage. Source: BHPS.

Occupational code	Occupational sector	Likelihood Ratio	Occupation 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 Occupations 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 Occupations in Agriculture, Forestry and Fishing	16.96	0.0108
92	Other Occupations in Construction	11.78	0.0096
94	Other Occupations in Communication	0.42	0.0080
95	Other Occupations in Sales and Services	9.22	0.0336
99	Other Occupations NEC	0.45	0.0120
	Average (unweighted)	5.69	
	Average (weighted)	4.71	

Table 14. Occupational Persistence (Likelihood Ratios), 2–digit level. In the table we report the likelihood ratios of the occupations that employ at least 0.5% of the workforce, due to the limited size of the sample. Averages are taken with respect to all occupations, including the ones not reported in the table.

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 15. Occupational Persistence Indexes, by Father's Income (Likelihood Ratios).

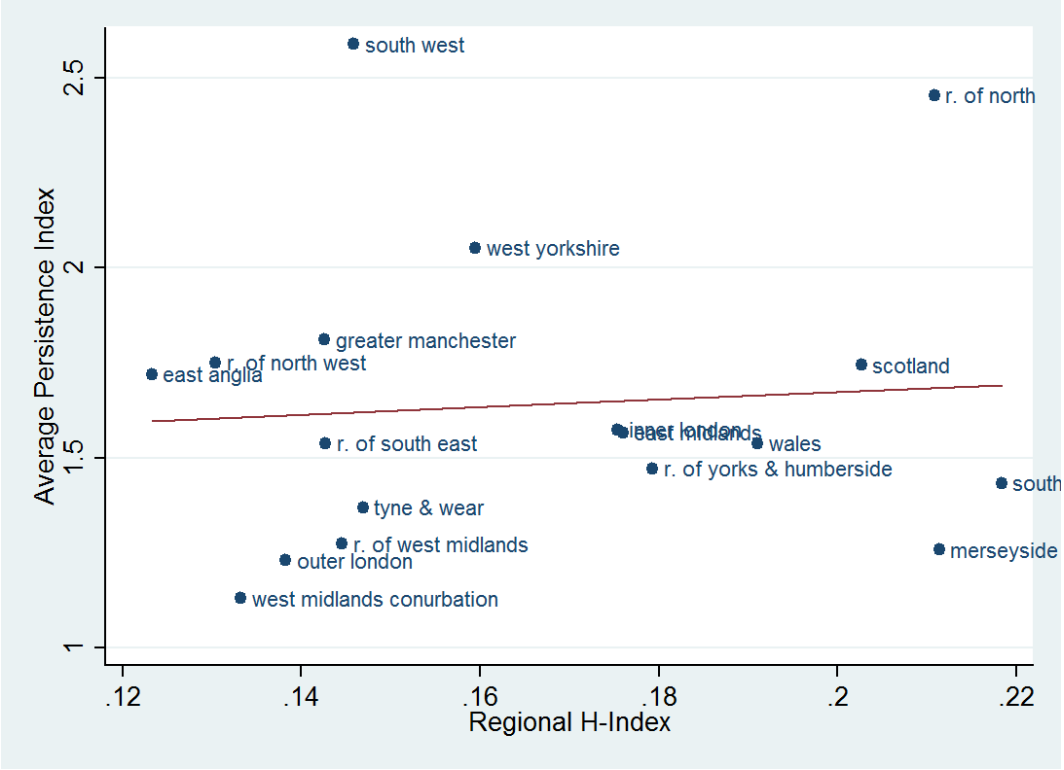

Figure 5. Plot of Average Persistence Index (weighted average of occupation–specific likelihood ratios, by region) vs. Herfindahl index of occupations (at the 1–digit level) . Source: BHPS.

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 16. Occupational Persistence Indexes, by Offspring's Age Group (Likelihood Ratios).

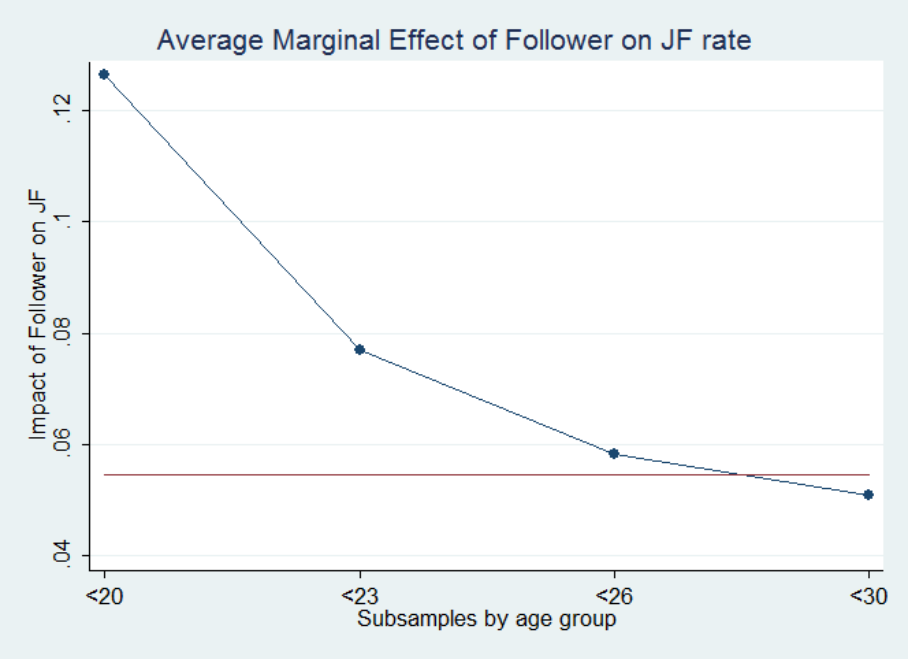

Figure 6. Average marginal effect of $\pi_{i,t}$ (coefficient of Column 1 of Table 4), by age group. The red line is the average marginal effect for the entire sample. Source: BHPS.

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 17. 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, 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, occupation of search/employment. In Panel A, we exclude from the estimation the spells of self–employment. In Panel B, we exclude all the workers who report having been self–employed at least once in their life.

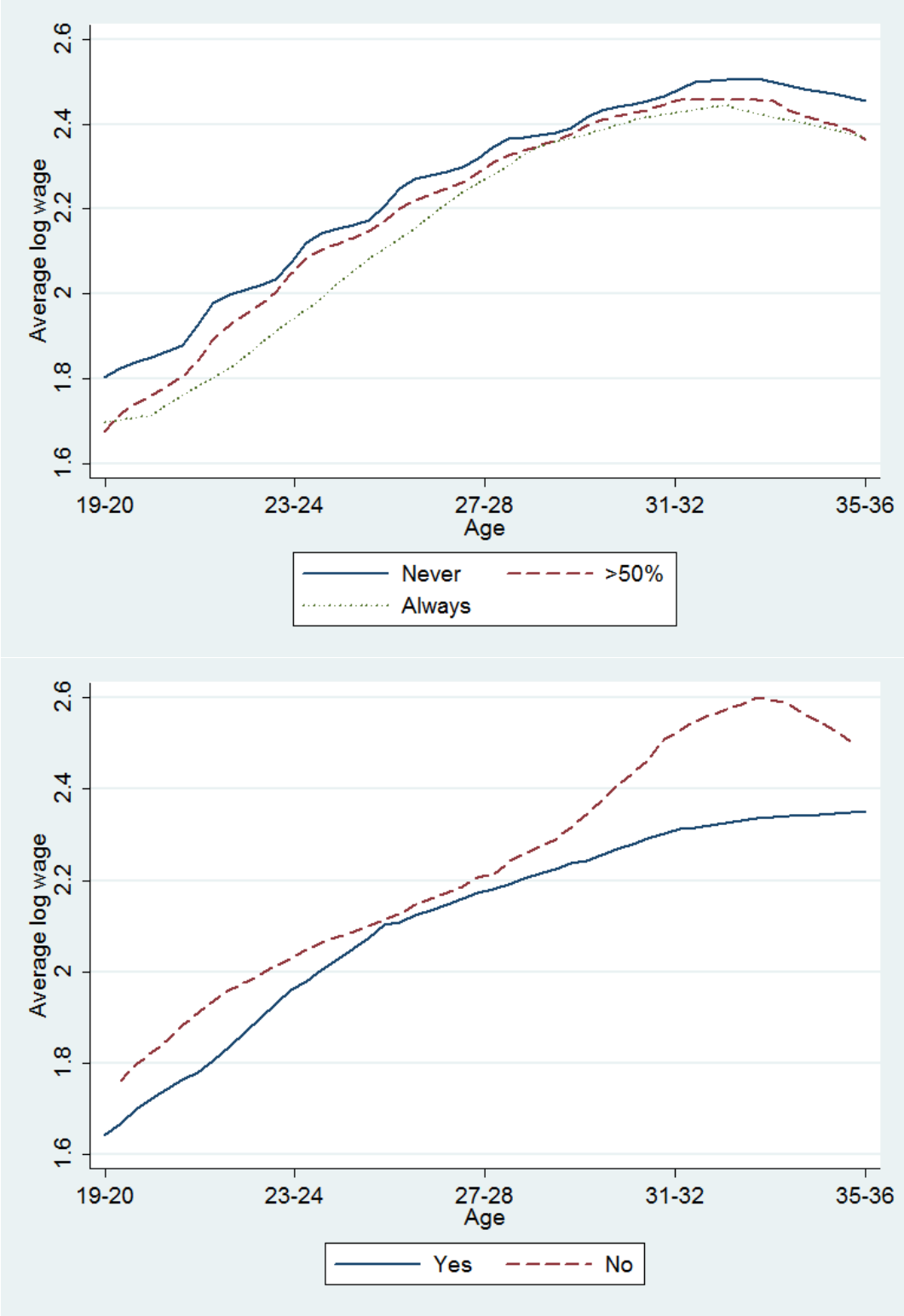

Figure 7. Wage profiles by proportion of job spell with the father in the same occupation. In the upper graph, groups are defined with respect to the entire job spell length. In the lower graph, groups are defined with respect to the start of the job spell. Source: BHPS.

Dependent Variable: Log Hourly Wage

	(1) POLS	(2) RE	(3) POLS	(4) RE	(5) FE
Panel A					
Share of time in same occ. as father (q_i)	-0.134***	-0.155***			
	(0.016)	(0.031)			
Father in same occupation ($\pi_{i,t}$)			-0.0730***	-0.0316**	-0.00698
			(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 same occ. as father (q_i)	-0.0967***	-0.109***			
	(0.015)	(0.029)			
Father in same occupation ($\pi_{i,t}$)			-0.0504***	-0.0190	-0.00349
			(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 18. 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, model 5 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education and occupation, a second order degree occupational tenure and potential labor market experience polynomial, firm size, region of residence, smoking behavior, marital status, ethnicity, 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.