The Costs of Occupational Mobility: An Aggregate Analysis

Guido Matias Cortes*

University of Manchester

Giovanni Gallipoli†

University of British Columbia

October 11, 2016

Abstract

We estimate the costs of occupational mobility and quantify the relative importance of differences in task content as a component of total mobility costs. We use a novel approach based on a model of occupational choice which delivers a gravity equation linking worker flows to occupation characteristics and transition costs. Using data from the Current Population Survey and the Dictionary of Occupational Titles we find that task-specific costs account for no more than 15% of the total transition cost across most occupation pairs. Transition costs vary widely across occupations and, while increasing with the dissimilarity in the mix of tasks performed, are mostly accounted for by task-independent occupation-specific factors. The fraction of transition costs that can be attributed to task-related variables appears fairly stable over the 1994-2013 period.

JEL Codes: J62, J24.

Keywords: Occupational Mobility; Tasks; Worker Flows; Mobility Costs; Gravity Model.

---

*Department of Economics, University of Manchester, Arthur Lewis Building, Oxford Road, Manchester, M13 9PL, UK. E-mail: matias.cortes@manchester.ac.uk.

Financial support from CLSRN and SSHRC is gratefully acknowledged. We thank Dirk Krueger and four anonymous referees for extensive comments that significantly improved the paper. We are grateful to Christopher Nekarda for facilitating access to the matched CPS data and to Todd Schoellman for generously sharing his occupation licensing database. We also thank Nicole Fortin, David Green, Keith Head, Thomas Lemieux, Vadim Marmer, John Ries, Uta Schoenberg, Henry Siu, Tomasz Swiecki, Yaniv Yedid-Levi, and seminar participants at various universities and workshops for insightful comments and suggestions. Carlos Sanchez and Lijing (Emily) Wang provided valuable research assistance at different stages of this project.

†Vancouver School of Economics, University of British Columbia, 6000 Iona Drive, Vancouver, BC V6T 1L4, Canada. E-mail: gallipol@mail.ubc.ca

1 Introduction

Several contributions to the human capital literature have analyzed the costs associated with different types of employment transitions. Topel (1991) provides evidence that a typical male worker in the United States with 10 years of job tenure loses 25% of his wage if his job ends exogenously. Other papers have analyzed the extent to which human capital is transferable across jobs. Neal (1995) and Parent (2000) argue that an important component of human capital is industry-specific and, therefore, only lost when a worker switches to a different industry. Kambourov and Manovskii (2009b) and Sullivan (2010) find evidence that a major component of human capital is occupation-specific. Meanwhile, Pavan (2011) presents evidence that tenure in a given career (which is empirically identified as a combination of industry and occupation), and tenure within a given firm are both important sources of wage growth over the life cycle.

In this paper we quantify the costs of occupational mobility and the relative importance of differences in task content as a component of these costs. As argued by Lazear (2009), Poletaev and Robinson (2008), Gathmann and Schönberg (2010), and Yamaguchi (2012), occupational transitions differ vastly in the extent of task switching that they entail. In some cases a worker may completely change careers, while other transitions involve only a minor adjustment in the mix of tasks performed. If human capital built in an occupation is task-specific, it should be partially transferable to occupations in which a similar set of tasks is performed. This paper develops a framework which allows for heterogeneity in the costs of transiting between occupation pairs according to the extent of overlap in their task content. We use this framework to estimate the size and the composition of occupational mobility costs.

The key innovation in this paper is the approach that we take in estimating transition costs. Specifically, we develop an occupational choice model inspired by the gravity framework typically used in the trade literature. In that literature, the interest is in estimating barriers to trade using data on flows of goods across countries, and proxies for trade costs that include geographical distance and whether the countries share a common border or a common language, among others. We show that this approach may be adapted to identify the costs associated with occupational mobility by using data on worker flows across occupations and proxies for mobility costs based on task data.

Previous studies of the costs of job transitions have relied primarily on wage data, particularly on the wage changes experienced by displaced workers when transiting into employment in a new firm, occupation or industry. Using this empirical approach, Gathmann and Schönberg (2010) and Poletaev and Robinson (2008) find evidence of the importance of task-specific human capital. A number of recent contributions also use wage data in combination with worker flow data to identify transition costs. Artuc et al. (2010) and Artuc and McLaren

(2015) estimate transition costs across industries and very broad occupation groups using the responsiveness of worker flows to inter-sectoral wage differences. Herz and Van Rens (2015) examine the relationship between job finding rates and wages across industries and locations to infer transition costs. Lalé (2015) identifies transition costs by linking the volatility of productivity shocks and wages to net occupational mobility.

The identification approach proposed in this paper is primarily based on worker flow data. By focusing on flows we effectively use information from occupation switchers as well as occupation stayers for identification. The approach is fairly simple and allows for worker heterogeneity in match-specific draws and for shocks to occupation-specific payoffs. A key advantage of our approach, in contrast to methods that rely directly on observed wages for identification of transition costs, is that we are able to adopt a broad notion of the costs and benefits of occupation transitions, entailing both pecuniary and non-pecuniary returns. Our procedure delivers identification of both the level and the change over time in occupational mobility costs.

Relating worker flows across occupations to the degree of skill transferability was an idea originally suggested by Shaw (1984).[1] We explicitly test whether worker flows are related to observable measures of task content and determine how much of the total cost of changing occupations can be attributed to task-related measures. Understanding the costs associated with the reallocation of workers across tasks is particularly relevant in light of recent literature on the polarization of the labor market, which suggests that technological change has altered the demand for particular tasks (e.g. Autor et al., 2003, 2006; Goos and Manning, 2007), and induced worker reallocation (Autor and Dorn, 2009; Cortes, 2016).

The theoretical setting which underpins our empirical analysis involves a partial equilibrium occupational choice model with perfect information, similar to the static framework used in a different context by Eaton and Kortum (2002). There is a continuum of workers, who differ in terms of observable characteristics, including the occupation which they start the period in. Workers make match-specific productivity (or match quality) draws from a set of extreme value distributions corresponding to each potential occupation. The extreme value assumption may be justified by thinking of workers as receiving a large set of offers from different employers within each occupation, and only considering the highest offer in each occupation. The distribution of highest offers (maxima) across employers within each occupation converges to an extreme value distribution.

Once workers observe their draws, they decide which occupation to work in during the period. There are costs to switching occupations, which depend on the particular occupation that the worker starts in, and the particular occupation that she considers switching to. Given the match-specific productivity draws and taking into account the costs and benefits

[1]Worker flows are also exploited in recent work by Sorkin (2015) to identify the importance of compensating differentials and rents for earnings differentials between firms.

of mobility, the worker chooses the occupation where she will receive the highest payoff. Workers' optimal switching decisions and the properties of the extreme value distribution lead to a gravity-type equation, which relates the flow of workers between any occupation pair to a set of occupation-specific characteristics and to the cost of switching.

In this environment selection into occupations hinges on idiosyncratic match-quality draws.[2] This occupational choice model has a long tradition in labor economics: for example, in his original taxonomy of earnings functions, Willis (1986) describes it as one of the classical models of selection.[3] The main disadvantage of this framework is that it is not suitable to study how workers progress through hierarchically ranked occupations, commonly referred to as job ladders. However, characterizing job ladders is not an objective of our study, as we instead focus on measuring different layers of costs for *all possible* transitions. Various papers analyze the relevance of job ladders or focus on selection patterns into hierarchically ranked occupations (see for example Gibbons, Katz, Lemieux, and Parent, 2005; Groes, Kircher, and Manovskii, 2015; Yamaguchi, 2012). Our contribution is complementary to, but distinct from, that literature. The match-quality model provides a parsimonious method to identify occupational mobility costs between any occupation pair, and offers a natural way to quantify the relative importance of the task content of occupations as one component of these costs. Crucially, this simplicity allows us to derive a concise estimable expression without imposing strong assumptions about the life-cycle evolution of wages.

To estimate the gravity equation obtained from the model we use variables that are related to the cost of an occupational switch. Following a growing literature, we characterize occupations through a vector of task characteristics (e.g. Autor, Levy, and Murnane, 2003; Ingram and Neumann, 2006; Gathmann and Schönberg, 2010; Poletaev and Robinson, 2008; Yamaguchi, 2012) using data from the Dictionary of Occupational Titles (DOT). We then construct a measure of distance between occupation pairs, which captures the degree of dissimilarity in the mix of tasks performed in the two occupations.[4] If a considerable share of human capital is lost when a worker experiences a dramatic change in the set of tasks she performs, the costs of occupational mobility should be increasing in task distance. We also allow for a fixed cost of switching across occupations that belong to different major task groups (non-routine cognitive, routine cognitive, routine manual or non-routine manual), as there may be costs associated with these switches in excess of what is captured by the dis-

[2]In addition to the match-quality shocks, our framework also allows for general occupation-level shocks to affect selection into occupations (as in Kambourov and Manovskii, 2009a).

[3]More recently, Gottschalk et al. (2015) consider a setup with an analogous sorting mechanism (the "independent productivity shocks model" in their paper) in order to obtain bounds on the changes over time in task prices. See also Helpman et al. (2010) for a model where workers sort into firms based on idiosyncratic match-quality draws.

[4]Yamaguchi (2012) develops a framework that distinguishes between task content and individual skills, and finds that "the derived policy function for occupational choice suggests that observed tasks can be interpreted as a noisy signal of unobserved skills."

tance measure. Finally, we allow for occupation-specific entry costs, which are independent of task content. These costs partly reflect institutional barriers faced by potential entrants to an occupation, such as qualification credentials, professional training, licensing and union membership requirements.

We estimate the gravity equation using data on monthly worker flows across 2-digit occupations from the matched monthly Current Population Survey (CPS) from 1994 to 2013. This is a period during which the CPS employed dependent coding techniques, which have been shown to reduce the amount of coding error in occupational transitions (Moscarini and Thomsson, 2007; Kambourov and Manovskii, 2013).

We find that task distance is a significant component of the cost of switching occupations, suggesting a role for task-specific human capital. An increase of one standard deviation in task distance within an occupation pair is estimated to increase the cost of switching by approximately 14%. This would translate into a 43% fall in the ratio of switchers to stayers in the source occupation. If the destination is in a different major task group, the cost is increased further, between 23 and 68 percentage points, depending on the type of transition. In spite of the significant role of task content, we find that the fraction of transition costs attributable to these variables is relatively small. For the median occupation pair, task-related costs account for only around 6% of total costs. Task-related costs account for more than 15% of total costs only for 5% of our observations. The remainder is accounted for by the task-independent occupational entry costs. These task-independent entry costs are shown to be correlated with a number of empirical measures of occupational access costs, including the amount of specific vocational preparation required for average performance in the occupation and the number of states in which the occupation is subject to licensing requirements.

Through a set of counterfactual experiments we estimate the hypothetical increases in mobility rates that would be observed if transition costs were reduced. For the median occupation in our sample, we find that the hypothetical increase in mobility if task-related costs were removed is approximately 7.5 percentage points. Given that monthly occupational mobility rates in the sample range between 2% and 10%, this increase is considerable. However, this increase represents only around 11% of the total increase in mobility that would be observed if we also reduced other costs – namely, the task-independent occupational entry costs – to the lowest observed value in the sample. This implies that there is a substantial amount of heterogeneity in task-independent entry costs across occupations, and confirms our finding that the majority of the costs of occupational mobility are attributable to task-independent costs.

Our estimation allows all of the components of the transition costs to vary over time. However, we find that the fraction of transition costs that can be attributed to task-related variables remains stable over time, in the range of 10 to 14%.

We verify the robustness of our results in several ways. We use alternative task dimensions from the DOT and from its successor, O*Net. We restrict the analysis to younger workers for whom occupational mobility rates are higher and for whom our model assumptions about the occupational choice process may be more realistic (Neal, 1999; Gervais, Jaimovich, Siu, and Yedid-Levi, 2016). We also verify robustness by performing our estimation using only college-educated workers. Finally, we confirm that our results go through under a very wide range of assumptions about the dispersion of match quality draws.

Our findings regarding the large magnitude and heterogeneity of the transition costs across occupations are in line with the results of Artuc, Chaudhuri, and McLaren (2010) and Dix-Carneiro (2014) who, using different identification strategies, find that both the mean and the standard deviation of workers' moving costs across industries are high, amounting to several times average income.[5]

The rest of the paper is organized as follows. Section 2 describes how a gravity model similar to the one developed by Eaton and Kortum (2002) can be used to study flows of workers across occupations and the costs of occupational mobility. Section 3 describes the empirical strategy and data sources. Section 4 presents the findings of the paper. Section 5 discusses a number of robustness checks and extensions. Section 6 presents a general discussion of our results, while Section 7 concludes.

2 Model

The economic environment is a variant of that in Eaton and Kortum (2002), modified to account for flows of workers across occupations and the costs of occupational mobility. It involves a static partial equilibrium model with perfect information. All equations in this section hold at any period t, so for simplicity time subscripts are omitted.

2.1 Workers and Occupations

There is a continuum of workers of measure one indexed by i, and a finite set of occupations given by $j \in \{1, 2, ..., N\}$ with a large number of employers in each occupation. Workers differ in terms of observable characteristics and initial occupation. A worker's occupation at the beginning of the period is predetermined and indexed by k.

Workers select endogenously into occupations in order to maximize their utility payoff. The potential payoff in each occupation is individual-specific and can be interpreted as a total lifetime payoff which includes pecuniary benefits (i.e. wages), as well as non-pecuniary returns related to an individual's preference for each particular occupation. Switching occupations is costly, so if individual i selects into an occupation other than her current occupation k she faces an iceberg cost which is occupation-pair specific. This cost may be related to the

[5]Similarly, Artuc and McLaren (2015) document high switching costs in a model featuring five broad occupation categories.

pecuniary component of the payoff: for example, switchers may receive lower payoffs in their new occupation due to the fact that they need to learn a new set of tasks and are therefore less productive than incumbents. It may also be related to the non-pecuniary component of the payoff: for example, switchers may have to overcome certain institutional barriers in order to enter into a new occupation, some of which may not necessarily be reflected in their post-switching wages.

2.2 Payoffs

Let the potential payoff from selecting into occupation j for worker i whose initial occupation is k be denoted by $\phi_j(i|k)$. This payoff is given by:

$$\phi_j(i|k) = p_j f\left[X(i)\right]\left(\frac{z_j(i)}{d_{kj}}\right). \tag{1}$$

p_j is a single index subsuming those occupation features which affect all individuals in occupation j. It can be interpreted as a measure of the general attractiveness (in terms of utility payoff) of this occupation. $X(i)$ is a vector of individual characteristics, such as education and overall work experience, which reflect general human capital and change the returns for individual i in all potential occupations.[6] $z_j(i)$ is a match-quality shock reflecting how well worker i is matched with occupation j in terms of productivity and preferences. The process by which individuals draw the match-specific component of payoffs is described below. d_{kj} represents the cost of switching between the worker's initial occupation k and the potential occupation j, with $d_{kk} = 1$ (staying in the same occupation is costless) and $d_{kj} > 1$ for all $j \neq k$. Intuitively, this cost captures time and efficiency losses associated with learning and adapting to a different occupation.

$\phi_j(i|k)$ can also be interpreted as the potential wage paid to individual i if she selects into occupation j. If each occupation j produces a different final good with price given by p_j and labor is the only input in production, then Equation (1) would hold as long as wages are proportional to the marginal product of labor. The logarithm of Equation (1) is:

$$\ln \phi_j(i|k) = \ln p_j + \ln f\left[X(i)\right] - \ln d_{kj} + \ln z_j(i) \tag{2}$$

which is similar to wage specifications commonly used in the empirical literature, with $\ln p_j$ representing an occupation wage premium and $\ln f\left[X(i)\right]$ the return to a set of observable characteristics. Here the equation also includes the switching cost $\ln d_{kj}$, and has a match quality term that is extreme-value distributed as described below. It is important to emphasize that Equation (2) only represents *potential* rather than observed log wages, so it would not

[6] Appendix A describes how the model can be extended to include occupation-specific human capital. This extension does not affect the empirical strategy described below.

be possible to directly estimate it.

Both interpretations of $\phi_j(i|k)$ – as a utility payoff or as a wage – are compatible with our empirical strategy, as our estimation relies on worker flows across occupations rather than on observed wage or payoff data. We prefer the interpretation of $\phi_j(i|k)$ as a total utility payoff because it captures the fact that occupational choices are based on both pecuniary and non-pecuniary factors, and on long-run considerations that are not necessarily reflected in current wages.

2.3 Match Quality Draws

For each occupation j, individuals draw $z_j(i)$ from a Fréchet distribution.[7] One can think of individuals as receiving job offers from several potential employers in each occupation. The only offer the individuals will consider will be the best offer in each occupation (the one that offers the highest match quality). Thus, the set of 'relevant' offers for each occupation is the collection of maxima across firms for each individual. The distribution of the maxima of a set of draws can converge to one of only three distributions, one of which is Fréchet (type II extreme-value).[8]

The Fréchet distribution has a CDF given by:

$$z_j \sim F_j(z) = e^{-T_j z^{-\theta}} \tag{3}$$

The occupation-specific parameter T_j governs the location of the distribution. Match quality draws are on average higher in occupations with a larger T_j. The parameter θ, which is common across occupations, is related to the dispersion of the shocks, with a larger θ implying less variability.

Individuals sample occupation match-qualities at the beginning of the period, drawing a value for each occupation including their current one. They then compare potential payoffs, based on realized draws and transition costs d_{kj}, and choose whether to switch to a different occupation.[9] Each draw is independent from all others and corresponds to a guaranteed job offer from an employer in an occupation, so the individual faces no uncertainty when choosing an occupation. Section 2.5 discusses a number of alternative specifications of the model, where we allow the distribution of match-quality draws to differ according to workers' current occupation, and we discuss how the model can be re-cast in order to allow for search frictions.

[7] A similar setup is used by Hsieh, Hurst, Jones, and Klenow (2013).

[8] See Eaton and Kortum (2002), footnote 14, and references therein.

[9] Individuals make occupational choices after observing their match quality draws and consider the costs and benefits of potential transitions. Hence, random assignment of workers to occupations would not result in the same allocation of workers nor in the same equilibrium level of output.

2.4 Flows of Workers Across Occupations

For individual i (who starts in occupation k), the probability that her payoff in occupation j is above some level ϕ is given by:

$$\begin{aligned} Pr\left[\phi_j(i|k) > \phi\right] &= 1 - F_j\left(\frac{\phi d_{kj}}{p_j f\left[X(i)\right]}\right) \\ &= 1 - e^{-T_j d_{kj}^{-\theta}(p_j f[X(i)])^{\theta}\phi^{-\theta}} \end{aligned} \tag{4}$$

The probability that individual i obtains a payoff below ϕ in every occupation *other than* j is:

$$\begin{aligned} Pr\left[\phi_s(i|k) \le \phi, \forall s \neq j\right] &= \prod_{s\neq j} F_s\left(\frac{\phi d_{ks}}{p_s f\left[X(i)\right]}\right) \\ &= \prod_{s\neq j} e^{-T_s d_{ks}^{-\theta}(p_s f[X(i)])^{\theta}\phi^{-\theta}} \end{aligned} \tag{5}$$

Individual i will optimally choose to switch to occupation j, given her current occupation k, if j offers her the highest potential payoff among all possible occupations. The probability that this happens is denoted by $\pi_{kj}(i)$ and is given by:

$$\begin{aligned} \pi_{kj}(i) &\equiv Pr\left[\phi_j(i|k) \ge \max_s \left\{\phi_s(i|k)\right\}\right] \\ &= \int_0^{\infty} Pr\left[\phi_s(i|k) \le \phi, \forall s \neq j\right] \cdot dPr\left[\phi_j(i|k) \le \phi\right] \\ &= \frac{T_j d_{kj}^{-\theta} p_j^{\theta}}{\sum_{s=1}^{N} T_s d_{ks}^{-\theta} p_s^{\theta}}. \end{aligned} \tag{6}$$

Intuitively, j will be the best choice for individual i whenever j offers her a payoff ϕ while all other occupations offer her a payoff below ϕ. Integrating this over all possible values of ϕ gives the probability that i switches from k to j, $\pi_{kj}(i)$. This allows for the possibility that $j = k$, i.e. the optimal choice may involve staying in the current occupation. This probability is not individual-specific, so we can omit the i index.

Taking the ratio of π_{kj} and π_{kk}, based on Equation (6), and using the fact that $d_{kk} = 1$, we obtain

$$\frac{\pi_{kj}}{\pi_{kk}} = \frac{T_j d_{kj}^{-\theta} p_j^{\theta}}{T_k p_k^{\theta}}. \tag{7}$$

Or in logarithms:

$$\ln \frac{\pi_{kj}}{\pi_{kk}} = \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k - \theta \ln d_{kj} \tag{8}$$

With a large number of individuals in each occupation making independent draws from the match quality distribution, π_{kj} will be equal to the fraction of k-workers who switch to j, that is:

$$\pi_{kj} = \frac{sw_{kj}}{N_k} \tag{9}$$

where sw_{kj} is the total number of switchers from k to j and N_k is the size of occupation k (at the start of the period).

Therefore, Equation (8) can be rewritten in terms of worker flows, leading to a gravity-type equation:

$$\ln \left(\frac{sw_{kj}}{sw_{kk}} \right) = \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k - \theta \ln d_{kj} \tag{10}$$

This is the key equation of the model. It relates the flows of workers between occupations to a set of occupation-specific characteristics (T_j, T_k, p_j, p_k), and to the cost of switching (d_{kj}). In the empirical implementation, all of the variables in the equation will be allowed to vary period by period. Note that there are no individual-specific variables such as payoffs or wages in this equation.

2.5 Variations of the Benchmark Model

Equation (10) is consistent with various generalizations of the model. For example, the baseline model assumes that match-quality draws in destination occupation j are independent of source occupation k. However, it is conceivable that moving from certain occupations may be generally associated with higher (or lower) payoffs. To allow for this possibility, the distribution of match-quality draws can be re-written as

$$z_j \sim F_j(z|k) = e^{-\eta_k T_j z^{-\theta}}. \tag{11}$$

Under this specification, starting the period in occupation k offers an absolute advantage (or disadvantage) in all potential occupations j. This modification has no bearing on the derivation of Equation (10), as the parameter η_k shifts the distribution of draws for workers who start in occupation k but does not alter the relative probability of transiting to a particular occupation j.

An alternative way to recast the model is to posit no iceberg costs of switching but rather assume that the match-quality distribution in occupation j, for workers currently in occupation k, is:

$$z_j \sim F_j(z|k) = e^{T_j d_{kj}^{-\theta} z^{-\theta}}$$

This setup is empirically equivalent to the one in the baseline model. In other words, the transition cost d_{kj} may be modeled as impacting the quality of the offers received in occupation j by workers currently in occupation k, or as reducing their payoffs in occupation j conditional on an offer quality. These alternative specifications result in the same gravity representation.

The baseline model assumes that there are no search frictions, so that workers obtain match quality draws for all potential occupations j. One could obtain a specification analogous to Equation (10) by instead positing that the *probability* of receiving an offer from a potential occupation j, for workers currently in occupation k, depends on characteristics of the source and destination occupation and is proportional to d_{kj}. In this alternative context worker flows across occupations would still be a function of the characteristics of the source and destination occupation and of d_{kj}; however, the driving force restricting flows between occupations k and j would be a scarcity of job offers, rather than a reduction in payoffs due to iceberg costs. Our objective in this paper is not to distinguish whether the barriers to mobility operate through a lack of job offers or through iceberg costs, but rather to estimate the magnitude of these barriers and assess the relative importance of task content as a component of overall transition costs.

3 Data and Empirical Implementation

Our objective is to estimate Equation (10) and quantify the costs of switching between different occupation pairs, d_{kj}, and the factors that affect these costs. One such factor is the 'task distance' between occupations k and j. As suggested by Gathmann and Schönberg (2010), Poletaev and Robinson (2008) and Robinson (2011), if human capital is task-specific, it should be partly transferable across occupation pairs in which a similar mix of tasks is performed. The cost of switching occupations should therefore be increasing in the degree of dissimilarity, or 'distance', in the task content of the two occupations.[10]

To construct a measure of task distance, we follow previous literature in characterizing occupations through a vector of skill or task characteristics (e.g. Autor et al., 2003; Ingram and Neumann, 2006; Poletaev and Robinson, 2008; Peri and Sparber, 2009). We do this using data from the Revised 4th Edition of the Dictionary of Occupational Titles (ICPSR, 1991). The DOT provides precise measures of the different abilities that are required in different occupations, as well as the different work activities performed by job incumbents. The dimensions along which the DOT dataset characterizes occupations include complexity of work, General Education Development (GED), specific vocational preparation requirements, aptitudes, temperaments and physical demands, among others (ICPSR, 1981). The choice of the relevant dimensions to characterize occupations and construct a distance measure is somewhat arbitrary. We choose the three GED variables and the eleven aptitudes from the 1991

[10] Task distance here parallels the traditional use of geographic distance in gravity models of trade.

DOT as the relevant dimensions for our baseline measure, and test the robustness to different choices in Section 5.1. Table 1 provides examples of the DOT task vectors for four particular occupations.[11]

Following Gathmann and Schönberg (2010), we construct a distance measure across occupation pairs based on angular separation. The distance measure reflects the degree of dissimilarity in the mix of tasks performed in the two occupations, and can be interpreted as an 'intensive margin' description of an occupational transition.[12] Specifically, let x_k^a be the importance level of dimension a (one of the dimensions described above) in occupation k, and x_j^a the analogous measure for occupation j. The angular separation between the task vectors in the two occupations is given by:

$$AngSep_{kj} = \frac{\sum_{a=1}^{A}\left(x_k^a \times x_j^a\right)}{\left[\sum_{a=1}^{A}(x_k^a)^2 \times \sum_{a=1}^{A}(x_j^a)^2\right]^{1/2}} \tag{12}$$

where A is the total number of dimensions being considered. $AngSep_{kj}$ ranges between -1 and 1, and is increasing in the degree of overlap between the two vectors. We transform this to a distance measure $dist_{kj}$ which ranges between 0 and 1 and is increasing in dissimilarity:

$$dist_{kj} = (1/2)\left(1 - AngSep_{kj}\right) \tag{13}$$

In addition to the task distance between occupations, we consider the possibility that there are costs for switching between major task groups. Following the literature (e.g. Acemoglu and Autor, 2011), we group occupations into four broad task groups: non-routine cognitive, routine cognitive, routine manual, and non-routine manual. Appendix Table A.1 provides details on the occupations included in each of these broad groups. We allow for a direct cost of switching between task groups, and we let this cost differ by destination group. Specifically, we define four dummy variables, λ_{kj}^{NC}, λ_{kj}^{RC}, λ_{kj}^{RM} and λ_{kj}^{NM}, which are equal to one if occupations k and j are in different broad task groups, and destination occupation j is a non-routine cognitive, routine cognitive, routine manual, or non-routine manual occupation, respectively. These dummies reflect costs of switching between task groups that are not fully captured by the distance measure.[13]

We also allow for a destination effect m_j, which reflects general costs of accessing occu-

[11] Each DOT dimension is normalized to have mean zero and standard deviation one across the universe of standardized 3-digit occupations from Autor and Dorn (2013). More details are provided in Appendix C.

[12] As in most existing studies, our measure of task distance assumes that all task dimensions are weighted equally. In principle one could consider alternative approaches that explore the uneven role of different tasks for the portability of task-specific human capital. This analysis would however require additional assumptions about the transferability of task-related skills.

[13] In the robustness analysis of Section 5.1 we estimate a specification in which the costs of transiting between broad occupation groups depend also on the broad source group.

pation j that are not related to the task content of j or to the characteristics of the source occupation k. This access cost captures any occupation-specific (but not task-specific) factors that restrict mobility and are not transferable across occupations. For example they could reflect the fact that some occupations are difficult to access due to the use of hard-to-acquire skills, regardless of whether one comes from an occupation with a similar task mix. They may also include institutional barriers such as professional qualifications, specific training or other requirements.

Finally, the unobservable term ϵ_{kj} captures costs of occupational mobility between occupations k and j arising from any other factor. ϵ_{kj} is assumed to be an independently and identically distributed random variable with a standard normal distribution.

We therefore have the following specification for $\ln d_{kj}$:

$$\ln d_{kj} = \beta_1 dist_{kj} + \beta_2 \lambda_{kj}^{NC} + \beta_3 \lambda_{kj}^{RC} + \beta_4 \lambda_{kj}^{RM} + \beta_5 \lambda_{kj}^{NM} + m_j + \epsilon_{kj} \tag{14}$$

Substituting equation (14) into the gravity equation (10) and defining $S_k \equiv \ln T_k + \theta \ln p_k$ and $D_j \equiv S_j - \theta m_j$, we obtain:

$$\ln\left(\frac{sw_{kj}}{sw_{kk}}\right) = D_j - S_k - \theta\beta_1 dist_{kj} - \theta\beta_2\lambda_{kj}^{NC} - \theta\beta_3\lambda_{kj}^{RC} - \theta\beta_4\lambda_{kj}^{RM} - \theta\beta_5\lambda_{kj}^{NM} - \theta\epsilon_{kj} \tag{15}$$

This equation can be estimated period-by-period using data on worker flows across occupations. Given the assumptions above, the error term $\theta\epsilon_{kj}$ has a normal distribution and is orthogonal to all other regressors.[14]

3.1 Source and Destination Heterogeneity

S_k and D_j can be identified through source and destination occupation fixed effects, respectively. We impose no restrictions on the relationship between S_k and D_j in our estimation. An occupation k will be estimated to have a high S_k if outflows from that occupation are relatively low, all else equal (that is, conditional on task variables and destination fixed effects). In our setting we have that $S_k \equiv \ln T_k + \theta \ln p_k$, so a high S_k may either be due to average match quality in that occupation being high (a high T_k) or to the characteristics of that occupation being associated with high returns (a high p_k). S_k therefore will be high in occupations with high average utility payoffs.

Meanwhile, an occupation j is estimated to have a high D_j if inflows to that occupation are relatively high, conditional on task variables and source fixed effects. The model indicates

[14]This assumption does not rule out that occupations that are more desirable may be more costly to move into, say because of a limited number of training spaces for potential entrants. This would be reflected in the source and destination occupation fixed effects, which are discussed below.

that $D_j \equiv S_j - \theta m_j$, so a high D_j may either be due to occupation j having a high average utility payoff (high S_j), or to occupation j being relatively easy to switch into due to low entry costs (low m_j).

We can separately identify S_k and D_j for all occupations k and j, relative to a base occupation. Hence we can back out the implied value for θm_j and obtain an estimate of this cost for each occupation. The relationship $\theta m_j = S_j - D_j$ illustrates that an occupation exhibits lower estimated access costs whenever D_j is large relative to S_j. The intuition is straightforward: higher inflows into an occupation, all else equal, indicate that the occupation is easily accessible.

A caveat is in order, as this framework allows for an alternative interpretation. One could conceive of a scenario in which there are no costs of accessing occupations, but workers face an exit cost when leaving an occupation. For example, some occupations could generate lock-in effects due to stigma or sunk costs associated to accumulation of skills that are not valued in other jobs. In Appendix B we formalize this scenario and show that the difference $(S_j - D_j)$ still identifies the transition cost specific to occupation j, but this can be interpreted as an exit cost rather than an access cost. In the Appendix we also consider a model in which both access and exit costs are present, and show through a simple empirical decomposition that access costs appear to be relatively more important. Moreover, in Section 4.2 we show that our estimates of the occupation-specific costs are correlated with a number of direct measures of access costs, including the amount of specific vocational preparation required for average performance in the occupation and the number of states in which the occupation is subject to licensing requirements. Hence, in the remainder of the paper we focus on the interpretation based on access costs.

3.2 Measuring Flows

We measure flows of workers across occupations using matched monthly data from the Current Population Survey (CPS). The CPS is a monthly survey of approximately 50,000 households conducted by the Bureau of the Census for the Bureau of Labor Statistics. It is the main source of labor market statistics in the United States. We make use of the fact that the CPS has a rotating sample structure: households included in the CPS are sampled for four consecutive months, then leave the sample for eight months, and then return for another four months. Given this sampling structure, up to 75% of households are potentially matched across consecutive months. In practice, the fraction of households that can be matched is slightly lower (around 70%), primarily due to the fact that the CPS is an address-based survey, so households that move to a new address are not followed. Also, in certain months the CPS made changes to household identifiers, making it impossible to match households across these modifications.[15] Details about the algorithm used to match individuals across months can be found in Nekarda (2009).[16]

The main advantage of the CPS relative to other longitudinal datasets such as the Panel Study of Income Dynamics (PSID) is its large sample size and the fact that it is explicitly designed to be representative of the entire US population at each point in time. An additional benefit of the CPS is that, in January 1994, dependent coding techniques were introduced in order to reduce the interview burden and the possibility of occupation and industry misclassification.[17] These techniques substantially reduce the amount of spurious transitions across occupations (see Kambourov and Manovskii (2013) and Moscarini and Thomsson (2007)) and hence allow us to measure flows across occupations in a more reliable manner.[18]

To take advantage of the dependent coding techniques, we use data starting in 1994 for our analysis. The most recent period available in our dataset is December 2013. The sample is restricted to adults aged 18 to 65 who are not in farming occupations or in the military.

We perform our analysis at the 2-digit occupation code level. Finer occupational groupings (i.e. 3-digit codes) provide a level of aggregation that is too low to observe significant flows of workers across particular occupation pairs. Meanwhile, a higher level of aggregation (1-digit level) creates groups that are too coarse. Our 2-digit occupations are an aggregation of the harmonized occupation codes from Autor and Dorn (2013), which are adapted from Meyer and Osborne (2005). The full universe of occupations is listed in Appendix Table A.1.[19] Appendix C provides details on the procedure followed to merge the CPS and the DOT data.

Using 2-digit occupation codes for matched individuals who are observed across consecutive months, we construct monthly flows of workers across occupation pairs. The flow of switchers from occupation k to occupation j is defined as the number of respondents (weighted using CPS sample weights) who are employed in occupation k in month t and employed in occupation j in month $t+1$. To reduce noise, monthly flows are aggregated at an annual level. The annual flows constitute our measure of switchers sw_{kj}.

[15] This affects the period between June and September of 1995.

[16] We thank Christopher Nekarda for facilitating access to the matched CPS data. Additional variables were obtained from Flood et al. (2015).

[17] See `http://www.census.gov/cps/methodology/collecting.html`.

[18] We take advantage of dependent coding by focusing our analysis on occupation transitions that occur over consecutive months of employment. Naturally, other transitions may instead involve an intervening period of unemployment or inactivity. Unfortunately, dependent coding procedures are not applied when workers transition to employment from unemployment (or non-participation), nor when measuring flows over longer time horizons (which would effectively allow for intervening periods of non-employment). Section 5.4 provides a discussion of potential misclassification concerns, while Appendix G reports results obtained when considering transitions over 12-month horizons.

[19] Potential discontinuities induced in the occupational categories by changes in the occupation coding system used by the CPS in 2003 and 2011 are not of concern for our purposes given that we perform our estimation separately for each year and hence identification is obtained solely from variation across occupation pairs in worker flows at a given point in time.

4 Results

4.1 Gravity Equation Estimation

This section presents the results of the estimation of Equation (15) using CPS data on worker flows and the proxies for mobility costs described above. We estimate Equation (15) separately for each year in our sample. One issue that we need to address is the fact that there are occupation pairs for which flows are zero in specific years. This occurs for approximately 10% of our occupation pair-year observations.[20]

We deal with the issue of zero-flows in several ways. Column (1) of Table 2 shows the results from the estimation of Equation (15) for the year 2012 when observations with zero flows are dropped from the sample. In Column (2), we replace the zeros with the smallest value observed in the sample for the left-hand-side variable in Equation (15), and estimate the regression using OLS. Finally, in Column (3) the same replacement of zeros is done as in Column (2) but, following Eaton and Kortum (2001), a Tobit-style regression is estimated instead of using OLS.[21] All specifications include source and destination occupation fixed effects.

The table shows that the effect of task distance on worker flows is negative and significant, suggesting that task distance is an important component of the cost of occupational mobility. The estimate in Column (3) implies that, all else equal, a one standard deviation increase in distance leads to a 43% fall in the ratio of switchers to stayers. Meanwhile, the negative and significant coefficient estimates on the task switching dummies imply that switching into a different broad task group is costly, and more so when switching towards routine manual occupations.[22]

The estimated coefficients in the table provide information on the responsiveness of relative workers flows to the task variables. As Equation (15) shows, these coefficients correspond to $-\theta\beta$, implying that the findings discussed so far – such as the fact that higher distances are associated with lower flows of workers – could be driven either by high-distance switches being very costly (i.e. a high β) or by match quality shocks having a low level of dispersion (i.e. a high θ). We disentangle these two components in Section 4.3.

4.2 Occupation Access Costs

The estimated source and destination occupation fixed effects are also of interest. The omitted category, for both source and destination, is "Executives, administrators and managers"

[20]The issue of zeros in gravity equations is discussed in the trade literature; see Head and Mayer (2013) for an overview.

[21]See Head and Mayer (2013) for a discussion of the advantages of using this method.

[22]Appendix Figure A.1 illustrates goodness of fit for the specification in Column (3) by plotting the fitted values against the true values.

(occupation code 2).[23] Figure 1 plots the estimated source and destination fixed effects for each occupation j in the year 2012, obtained from the Tobit-type specification in Column (3) of Table 2. Note that the source fixed effects plotted on the x-axis correspond to our estimates of $-S_j$. Each circle represents a 2-digit occupation and its size. The numbers within each circle correspond to the 2-digit occupation codes reported in Appendix Table A.1.

The figure shows a strong correlation between estimated source and destination fixed effects. This implies that occupations that are larger sources of worker flows (those from which relatively more workers leave) also tend to be larger destinations (they attract relatively more workers). This pattern is indicative of heterogeneous levels of churn across occupations, with those in the top-right corner of the graph exhibiting higher levels of churn.

The model associates high churn occupations to lower access costs.[24] The intuition for this result – under our baseline interpretation and subject to the caveat discussed in Section 3.1 – is as follows: many people exit from these occupations, so it must be the case that they are relatively unattractive. Nonetheless, these occupations also experience large inflows of workers. Hence, it must also be the case that their access costs m_j are relatively low. An example of one such set of jobs is "Other administrative support occupations, including clerical" (occupation code 25).

At the opposite end, occupations in the bottom-left area of Figure 1 feature low churn and are interpreted as being relatively attractive due to the fact that few people leave. They also experience relatively low inflows, which indicates that it must be the case that they are difficult to access for outsiders (high m_j). Note that these occupations also tend to be populated by smaller numbers of workers. An example of one such set of jobs is "Health assessment and treating occupations" (occupation code 8).

These findings are consistent with results in Caines, Hoffmann, and Kambourov (2015) who, using German microdata, also find evidence of significant differences in access costs across occupations, with the most desirable ones having steep barriers to entry.

Below we show that the task-independent occupation access costs m_j are a very large component of total transition costs. To provide further economic content to the interpretation of these costs, we present evidence of their correlation with a number of direct measures of occupational access costs. Panel A of Figure 2 plots the estimated access cost θm_j for each occupation (averaged across all years in our sample) against a measure of Specific Vocational Preparation (SVP) obtained from the Dictionary of Occupational Titles. This measure reflects the amount of training, in months, required to learn the techniques, acquire the information, and develop the facility needed for average performance in a specific occupation. It captures job-specific training only. Higher levels of the SVP measure are correlated with higher oc-

[23] This implies the normalization $S_2 = 0$, and therefore $T_2 = p_2 = 1$, in each period. The constant is an estimate of $D_2 = -\theta m_2$, and is part of the destination fixed effects for all j.

[24] Recall that $\theta m_j = S_j - D_j$, so the sum of the estimated source fixed effect ($-S_j$) and the estimated destination fixed effects (D_j) for each occupation provides an estimate of $-\theta m_j$.

cupational access costs. Panel B of Figure 2 shows that the average estimated access costs also tend to increase with the average fraction of workers with a college degree in each 2-digit occupation. Thus, occupations with high estimated m_j may require more specific training, possibly not transferable to other occupations, even if they share a similar task content. Panel C provides some evidence of a positive relationship between access costs and rates of union membership by occupation. This suggests that unionization may also contribute to restricting access to an occupation, even for workers coming from jobs with similar task content.

Finally, and perhaps most notably, Panel D uses a measure of occupational licensing intensity. As discussed by Schoellman (2012), licensure is the strongest form of occupational access restriction, because workers are legally required to obtain a license from the government to practice their profession. To measure these requirements, Schoellman collects data on occupational licensing, at both federal and state level as of March 2010 using the CareerOneStop database, sponsored by the U.S. Department of Labor.[25] Although some occupations are federally licensed, most licensure is done at the state level. The database includes a list of 9,308 data points, each consisting of a license-occupation pair. Following Schoellman (2012) we consider several alternative ways to separate heavily licensed from more easily accessible occupations. Panel D of Figure 2 uses a measure which reports the number of states in which each 2-digit occupation is licensed.[26] The figure documents that occupations with higher estimated access costs tend to be licensed in a larger number of states.

A broad overview of the four panels in Figure 2 illustrates how access costs are associated with different institutional features in different occupations: for example, licensing and education restrictions appear to be important for access to medical, legal and university teaching occupations, whereas unionization is key for access to protective service jobs as well as office machine operation, mail distribution and school teaching jobs. Given this observation, and as further evidence for the interpretation of m_j as access costs, we combine the four empirical measures of access costs presented in Figure 2 into a single composite index. Panel A of Figure 3 uses a weighted average of the four measures (specific vocational preparation, fraction of workers with college degree, union membership rate, and state licensing requirements), where each measure is standardized to have mean zero and standard deviation one, and all four measures are equally weighted. We also perform a Principal Component Analysis (PCA) using the four empirical access cost measures and extract the first factor, which accounts for roughly 56% of the total variance of the four proxies. We then use the predicted score of the first factor in each occupation as a composite proxy of access costs. Results are plotted in

[25] We thank Todd Schoellman for generously sharing his licensing database. Licenses may cover a portion of an occupation or multiple occupations. The data are reported by state but not all states participate, and participating states may not report all licenses.

[26] We measure the number of states in which a 2-digit occupation is licensed as the employment-weighted average of the number of states in which each corresponding 3-digit occupation is licensed, where the employment weights are generated from the CPS data for the year 2010.

Panel B of Figure 3 and display a clear positive correlation with the access costs obtained from our estimation.

4.3 Match Quality Dispersion or Switching Costs? Estimating θ

The effect of the task variables on observed worker flows depends on the effect of these variables on transition costs but also on the dispersion of match quality shocks. As shown in Equation (15), the coefficients obtained from the estimation of the gravity equation correspond to $-\theta\beta$. To disentangle these two components, and to obtain an estimate of the costs of switching across occupations, it is necessary to estimate θ, which dictates the dispersion of match quality shocks. To this purpose we take advantage of the properties of the extreme value distribution, which lead to a simple relationship between the distribution of payoffs predicted by the model and the parameter of interest, θ.

The extreme value distribution in Equation (3) has mean $T_j^{1/\theta}\Gamma\left(1-1/\theta\right)$, where Γ is the Gamma function. Its logarithm has a Gumbel distribution, with standard deviation equal to $\pi/(\theta\sqrt{6})$.[27]

The (ex-post) payoff for an individual starting in occupation k, $\phi(i|k)$, is also drawn from an extreme value distribution. Specifically, the probability that an individual ends up with a payoff below or equal to ϕ is equal to the probability that her potential payoff in *all* possible occupations is below or equal to ϕ. That is,

$$\begin{aligned} Pr\left[\phi(i|k) \leq \phi\right] &= Pr\left[\phi_j(i|k) \leq \phi \ \forall j\right] \\ &= \prod_{j=1}^{N} F_j\left(\frac{\phi d_{kj}}{p_j f\left[X(i)\right]}\right) \\ &= e^{-\left(\sum_{j=1}^{N} T_j d_{kj}^{-\theta}(p_j f[X(i)])^{\theta}\right)\phi^{-\theta}} \end{aligned} \tag{16}$$

Equation (16) implies that the ex-post payoffs for the set of individuals starting in occupation k follow an extreme value distribution with mean $\left[\sum_{j=1}^{N} T_j d_{kj}^{-\theta}(p_j f\left[X(i)\right])^{\theta}\right]^{(1/\theta)}\Gamma\left(1-1/\theta\right)$. The standard deviation of the logarithm of these ex-post payoffs is $\pi/(\theta\sqrt{6})$.[28] Therefore, for any set of individuals starting in occupation k with common demographic characteristics $X(i)$, we have that $\sigma_k^x = \pi/(\theta\sqrt{6})$, where σ_k^x denotes the standard deviation of ex-post (log) payoffs within the group. This implies that θ can be estimated from the dispersion of payoffs σ_k^x.

[27] The Gumbel distribution has the CDF $F(x) = exp(-e^{-(x-\mu)/\beta})$, with standard deviation $(\pi\beta)/\sqrt{6}$. For the logarithm of the productivity draws from the extreme value distribution in Equation (3), we can write $Pr(\ln z_j(i) \leq z) = Pr(z_j(i) \leq e^z) = F_j(e^z) = exp(-T_j e^{-\theta z})$, which is a Gumbel distribution with standard deviation $\pi/(\theta\sqrt{6})$.

[28] From the previous footnote the standard deviation of the logarithm of productivity does not depend on T_j. For the distribution of log payoffs, T_j would be replaced by $\left(\sum_{j=1}^{N} T_j d_{kj}^{-\theta}(p_j f\left[X(i)\right])^{\theta}\right)$. The standard deviation remains independent of this multiplicative constant.

So far our analysis has relied exclusively on worker flow data and we have interpreted the payoff from an occupation as a present discounted value combining pecuniary and non-pecuniary returns. Unfortunately, while CPS data allows one to measure the dispersion of pecuniary payoffs (wages), there is no direct empirical counterpart to gauge the dispersion of the total utility payoff. The dispersion of current wages may be an imprecise proxy of the true dispersion of utility payoffs. Specifically, suppose that the payoff ϕ can be expressed as the product of the wage w and a factor φ that scales up current wages to the total lifetime utility payoff (pecuniary and non-pecuniary) in an occupation. We would then have that the variance of total log utility payoffs is given by:

$$Var(\ln \phi) = Var(\ln w + \ln \varphi) = Var(\ln w) + Var(\ln \varphi) + 2Cov(\ln w, \ln \varphi) \tag{17}$$

If w and φ are positively correlated (that is, workers with high current wage in an occupation also have high lifetime wages and/or high non-pecuniary returns in that occupation), or if these two components are negatively correlated but $[Var(\ln \varphi) + 2Cov(\ln w, \ln \varphi)] > 0$, then the variance of wages will underestimate the variance of total payoffs ϕ, leading to an overestimation of the value of θ. Only if the negative correlation between w and φ is such that $Var(\ln \varphi) + 2Cov(\ln w, \ln \varphi) < 0$ will the identification of θ based on the dispersion of wages underestimate the true value of θ.

Below we identify a benchmark value of θ based on wage dispersion. In Section 5.2 we gauge the robustness of our results to a wide range of estimates of θ, effectively allowing for different potential correlations between pecuniary and non-pecuniary rewards and between current and lifetime payoffs. It is worth emphasizing that wages have no role in the estimation of Equation (15). We use wage heterogeneity only to approximate the match quality dispersion parameter θ, and no information about wage *changes* is used to estimate mobility costs. Moreover, in Appendix D we outline an alternative approach to identify θ that does not use wage data at all. This approach yields estimates of θ that are similar to the ones based on wage dispersion. Finally, in Appendix E we present new evidence about the variance of non-pecuniary returns (approximated through a variety of job satisfaction measures from a different data source) and their correlation with wages. This evidence suggests that the omission of non-pecuniary returns implies a fairly small positive bias in the estimation of θ, with little or no change in results relative to our baseline parametrization.

Estimation. In what follows we present three alternative estimations of $\theta = \pi/(\sigma_k^x \sqrt{6})$ based on wage data.[29] First, we compute σ_k^x using the standard deviation of (ex-post) log wages from the entire set of individuals starting the period in occupation k (not conditional

[29]Wage data is available in the CPS for workers in the Outgoing Rotation Groups (fourth and eighth month in the sample). We follow the procedure in Lemieux (2006) to generate hourly wages and to trim extreme values of wages and adjust top-coded earnings.

on demographics). This is computed for each month in the sample using the distribution of wages in month t for workers with common occupation k in month $t-1$.

Next, to control for worker characteristics, we perform an estimation exercise based on residual wages. We first regress log wages on a flexible function of age and education.[30] We then compute the standard deviation of residual wages among all individuals starting the period in occupation k and use this to back out an implied value of θ.

Finally, we provide an estimate of θ based solely on data for young workers. As these individuals are at the start of their working life, there should be little heterogeneity in terms of life cycle shocks that affect wages. We compute estimates of $\theta = \pi/(\sigma_k^x\sqrt{6})$ based on the standard deviation of wages for people aged 25 to 30, by initial occupation and by gender.

Figure A.2 in the Appendix displays histograms of the estimated values of θ for each of these three methods.[31] Table 3 presents the corresponding summary statistics. The estimates of θ are higher when we condition on demographic characteristics. This is because the approach that does not condition on demographic characteristics overestimates σ_k^x and hence underestimates θ. The median estimate of θ based on the dispersion measures for young workers, in Column (3), is 3.23. Below we use this value as a baseline to compute implied mobility costs and perform counterfactual experiments. Section 5.2 gauges the sensitivity of baseline findings to alternative assumptions about the value of θ.

4.4 Mobility Costs

Table 4 uses the baseline value of $\theta = 3.23$ to measure the effect of different task variables on the iceberg transition cost d_{kj}. The first three columns of Table 4 show, for each of the specifications in Table 2, the estimated marginal effects $\widehat{\beta}$ of each of the variables on the logarithm of the transition cost ($\ln d_{kj}$). The next three columns compute the implied percentage effect on d_{kj} from a one standard deviation change in distance, and from a change from 0 to 1 for each of the task switching dummy variables. The results show that the impact of task distance on the cost of switching is substantial. For example, based on the coefficients from the Tobit-type specification, Column (6) shows that if distance increases by one standard deviation, the cost of switching occupations increases by approximately 14%, all else equal. Meanwhile, the switching cost is substantially increased if the switch involves a transition into a different task group. These additional costs range from 23% for transitions into routine cognitive occupations, to 68% for transitions into routine manual jobs.

[30] We include four education dummies, a full set of age dummies, and interactions of education dummies with a quartic in age. We also include month dummies to account for seasonality, and estimate separate regressions for each year and gender, thus flexibly allowing returns to education and age to vary between genders and over time.

[31] We use occupation-month bins with at least 100 observations for the first two approaches, and occupation-gender-month bins with at least 15 observations for the approach that considers only young workers.

Decomposing Mobility Costs

Next, we calculate the estimated transition cost d_{kj} for specific occupation pairs based on Equation (14).[32] We compute this in three steps. First, we calculate the cost solely attributable to task distance:

$$\ln d_{kj}^{dist} = \widehat{\beta}_1 dist_{kj}$$

We then add the cost attributable to switching between broad task groups, in order to obtain the total cost associated with the task variables:

$$\ln d_{kj}^{tasks} = \widehat{\beta}_1 dist_{kj} + \widehat{\beta}_2 \lambda_{kj}^{NC} + \widehat{\beta}_3 \lambda_{kj}^{RC} + \widehat{\beta}_4 \lambda_{kj}^{RM} + \widehat{\beta}_5 \lambda_{kj}^{NM}$$

Finally we quantify the full estimated transition cost, considering all costs including the fixed destination entry cost:

$$\ln d_{kj}^{all} = \widehat{\beta}_1 dist_{kj} + \widehat{\beta}_2 \lambda_{kj}^{NC} + \widehat{\beta}_3 \lambda_{kj}^{RC} + \widehat{\beta}_4 \lambda_{kj}^{RM} + \widehat{\beta}_5 \lambda_{kj}^{NM} + \widehat{m}_j$$

Table 5 shows estimates of these three layers of costs for a selected number of occupation pairs in year 2012. The top half of the table lists occupation pairs that exhibit the lowest overall transition costs, while the bottom half presents the occupation pairs for which we estimate the highest transition costs. All low-cost transitions have a low task distance and do not entail switching between broad task categories. They are also transitions into occupations with relatively low entry costs. Yet, even for such transitions estimated costs remain fairly large. Recall from Equation (1) in the model that d_{kj} is an iceberg cost which reduces the payoff to a worker who switches occupations. For example, the estimated cost of 1.043 associated with the task costs for transitions between "financial records processing occupations" and "other administrative support occupations, including clerical" implies that a switcher's payoff would be 4.3% higher if there were no costs associated with the task content of occupations (put differently, the utility payoff of an incumbent in the occupation with identical characteristics would be 4.3% higher). Overall, the payoff to a worker switching between these two occupations would be more than 3 times higher if all mobility costs were removed. Estimated switching costs are therefore substantial, even across occupations that see relatively high volumes of flows.

Transitions between occupation pairs at the bottom of Table 5 are the most costly. These transitions involve a high task distance, a transition into a different broad task group, and a transition into occupations with high task-independent entry costs (lawyers and judges). Estimated transition costs are in fact prohibitively high and we observe essentially no transitions between these occupations.

[32]We use the results from the Tobit-type specification in Column (3) of Table 2.

As a simple characterization of the relative importance of task variables, we compute the size of the iceberg cost associated with the task variables relative to the overall estimated iceberg cost. That is, $(d_{kj}^{dist}-1)/(d_{kj}^{all}-1)$ for the case of task distance, and $(d_{kj}^{tasks}-1)/(d_{kj}^{all}-1)$ for all task-related costs. Table 6 presents summary statistics for these ratios across all occupation pairs, and using all years in the sample. For the median observation, task-related costs account for approximately 6% of the total costs. Task-related costs account for more than 13% of all costs only for 1 in 10 occupation pair-year cells.[33]

Counterfactual Changes in Mobility Rates

An alternative route to measuring the magnitude of the estimated transition costs is to calculate counterfactual occupational mobility rates that would be observed if transition costs were reduced. Although the reduction in transition costs considered in this section may not always be realistic, nor necessarily desirable, these counterfactual exercises offer a simple way to gauge the relative importance of task-related barriers as compared to non-task related costs.

Column (1) of Table 7 shows the observed mobility rates towards other occupations for a number of 2-digit source occupations in year 2012 (that is, the total number of switchers as a fraction of the total number of workers in each occupation). The top half of the table includes occupations with the lowest observed outflows (between 2% and 4% of workers are observed to switch out at a monthly frequency), while the bottom half of the table includes the occupations with the highest outflows (over 5.5%). The final row shows aggregate mobility across 2-digit occupations. Column (2) presents fitted outflows, based on the estimation of the gravity equation.

Column (3) shows the estimated counterfactual mobility rates that would be observed if the switching costs associated with task distance were eliminated. As one would expect, the counterfactual increase in mobility relative to the fitted value in Column (2) is particularly large for occupations that are very remote, in the sense that they exhibit large task distances relative to most other occupations. An example of this type of remote occupation is "freight, stock and material handlers" where removing task distance alone increases occupational mobility rates from 8.3% to 11.8%.[34]

Column (4) displays the counterfactual mobility rates if one also removes costs associated with transitions across broad task categories. As expected, this induces further increases in mobility rates. Overall, for three quarters of the occupation-year cells in our sample,

[33] As an additional metric for the relative importance of the task variables, we note that the McFadden pseudo R-squared that would be obtained from the estimation in Column (3) of Table 2 if only the task variables are included is only 0.03. The corresponding R-squared for the full model is 0.16.

[34] The measure of task distance observed across the two most similar occupations in our sample is close to zero, so performing an alternative experiment where we reduce distance to the smallest value observed in the sample (instead of completely removing distance costs) yields essentially the same results.

counterfactual mobility rates increase by at least 5 percentage points (relative to the fitted values) when all task costs are removed. For 10% of the occupation-year cells, mobility rates increase by more than 12 percentage points. The increase for the median occupation in our sample is of 7.5 percentage points. This change is substantial, and approximately equal to the difference between the mobility rate of "helpers in construction and production occupations" (the occupation with the highest mobility rates) and that of "lawyers and judges" (the occupation with the lowest mobility rate). Meanwhile, the counterfactual increase in aggregate mobility is even larger, at just under 10 percentage points.

Column (5) calculates the counterfactual mobility rates that would be observed if, in addition to removing task-related costs, occupation-specific entry costs m_j are reduced to the lowest value observed in the sample. Clearly these task-independent entry costs are very important, as the counterfactual mobility rates in Column (5) are substantially larger than in Column (4).

The results in Column (5) provide a useful benchmark to assess the importance of task-related costs relative to general task-independent entry costs. Specifically, one can compare the increase in mobility rates that occurs when task-related costs are removed to the total increase that occurs when also task-independent costs are reduced to the lowest observed value. The main result from this exercise is presented in Table 8, which provides summary statistics for these relative changes using all occupation-year observations. For the median occupation, task-related costs only account for around 11% of the counterfactual mobility increase. Even at the top of the distribution – in occupations for which task content accounts for the biggest increase in counterfactual mobility – this fraction remains generally well below 25%.[35]

Columns (6) and (7) of Table 7 illustrate how our framework could be used to perform simple counterfactual experiments related to specific policy changes. As shown in Panels C and D of Figure 2, occupations with higher union membership and/or licensing rates tend to feature higher access costs. Given this observation we approximate the extent to which mobility rates would change following a reduction in either unionization or licensing rates. First, we compute the difference between average access costs among occupations with union membership rates above and below 9.8% (the median employment-weighted unionization rate across occupation-years). Column (6) explores a counterfactual where we reduce access costs

[35] An alternative counterfactual experiment in which all transition costs are completely eliminated yields mobility rates that are unreasonably high. A similar result applies in the trade literature: Counterfactual outcomes for a world with no trade costs imply implausibly higher levels of trade than what is observed in reality (Eaton and Kortum, 2002). In fact, in a world with no trade costs the share of a country's expenditure on its own goods would be proportional to its relative weight in the world economy. Analogously, here, the fraction of non-switchers in the counterfactual with no transition costs would be proportional to the source occupation's relative overall "attractiveness" within the universe of occupations, specifically: $\frac{T_k p_k^\theta}{\sum_j T_j p_j^\theta}$. Given a total of 37 occupations, this term would be equal to 0.027 if all occupations had the same T_j and p_j. In our sample, it ranges between 0.003 and 0.420.

for occupations with unionization rates above 9.8% by an amount equivalent to this difference. Results suggest that this change would increase aggregate mobility rates from 5% to 6%. Column (7) performs an analogous counterfactual experiment where we reduce access costs for occupations with above-median licensing rates. In this case, aggregate mobility rates increase to 7.2%. As one might expect, these changes are smaller than those obtained from counterfactuals in which task-related barriers are set to zero. However, they convey valuable information about the magnitude of the effects that could be realistically expected following viable policy changes. These results also suggest that, despite the fact that task-content accounts for a relatively small part of total transition costs, its effects on mobility are non-trivial when compared to those implied by realistic policy changes.

Changes Over Time

Given that we estimate Equation (15) separately for each year in our sample, we are able to assess the evolution of the estimated coefficients on the task variables over time. Figure 4 shows that the estimated coefficient on task distance has become smaller in absolute value over time. The difference between the estimated coefficient at the beginning and the end of the sample period is statistically significant. Under the assumption that the value of θ is constant, this result suggests that the marginal effect of task distance on transition costs has diminished over time.[36]

The estimated coefficients on the transitions across broad task groups do not display significant changes over time. They are presented in Appendix Figure A.3. The estimated source and destination fixed effects do not vary much over time either, suggesting that the ranking across occupations in terms of attractiveness and entry costs remains fairly stable over the 1994-2013 period.

We also verify whether the fraction of the transition costs that can be attributed to tasks varies over time, by performing exercises analogous to those in Tables 6 and 8 separately for each year. In spite of the reduction in the coefficient on task distance, we find that the fraction of transition costs that can be attributed to task-related variables remains stable over time, in the range of 5 to 6.5% for the median occupation according to the metric in Table 6, and between 10 and 14% according to the metric in Table 8.

[36] An alternative interpretation is that the dispersion of match quality has increased (i.e. θ has decreased). In fact, the observed increase in wage inequality suggests that this may be the case. If we estimate θ separately for each year in our sample we do observe a decrease in the estimated value of θ in the post-2000 period; however, the magnitude of the fall in the estimate of θ is not sufficient to account for the change in the coefficient on task distance, suggesting that there is a true decline in the marginal effect of task distance on transition costs over time.

5 Robustness

5.1 Alternative Measures of Task Content

To gauge the robustness of the results we consider a number of alternative measures for the construction of task distance. First, we add additional dimensions from the 1991 Dictionary of Occupational Titles to our task vector for the construction of the distance measure. We also consider alternative task characterizations which are based on more recent data from O*Net, the successor to the DOT.[37] Detailed results are reported in Appendix F. The fraction of the transition costs that can be attributed to task variables varies slightly when using alternative distance measures, but is no larger than 15% for the median occupation.

We also consider a specification where we allow for non-linear effects of task distance, and a specification in which transition costs between different broad task groups vary with both source and destination. As shown in Appendix F, the results from counterfactual experiments using these alternative specifications imply that task-related barriers account for around 10-13% of overall transition costs for the median occupation.

5.2 Alternative Values of θ

As discussed in Section 4.3, the dispersion of wages may differ from the dispersion of total payoffs depending on the correlation between current wages and other factors affecting lifetime payoffs, including the non-pecuniary component of match quality. To check the robustness of our results to alternative assumptions about the value of θ, and to further reduce reliance on the estimates obtained from wage data, Table 9 reproduces the results from Table 6 using a range of higher and lower values of θ. This implicitly allows for a wide range of assumptions about the variance of other components driving lifetime payoffs, and about the correlation of these components with current wages.

For values of θ that are below our benchmark estimate, the fraction of costs that is attributed to task-related variables relative to task-independent occupational entry costs is reduced further.[38] The fraction of costs attributed to task-related variables does not increase dramatically when higher values of θ are assumed: even when setting θ to a value as high as 8.87 (the highest value of θ obtained through the alternative estimation method described in Appendix D), task-related variables only account for around 13% of total costs for the median occupation pair.

It is also important to emphasize that the results from the counterfactual experiments

[37]We thank Nicole Fortin for sharing a crosswalk between O*Net occupation codes and occupation codes from the 1980/1990 Census occupation coding systems.

[38]From the discussion in Section 4.3, recall that our benchmark estimate of θ will be an upper bound estimate whenever $[Var(\ln\varphi) + 2Cov(\ln w, \ln\varphi)] > 0$. Moreover, Appendix E presents evidence that pecuniary and non-pecuniary returns are positively correlated.

presented in Table 8 are not affected in any way by the estimated value of θ. Our conclusion that the costs of occupational mobility are large and primarily driven by task-independent occupation entry costs are therefore robust to a wide range of assumptions about the value of the parameter θ.

5.3 Alternative Data Samples

Our next robustness check involves estimating the model for specific sub-samples of workers. The baseline model assumes that individuals receive productivity draws for all occupations and make switching decisions based on those draws. Results from previous literature, such as Neal (1999) and Gervais et al. (2016), suggest that this assumption is more likely to hold for younger workers who are at the start of their careers and still unsure about how well they will be matched to different types of jobs. In fact, the data shows that younger workers have higher rates of occupational mobility (Kambourov and Manovskii, 2008; Gervais et al., 2016).

Hence we verify the robustness of our results by estimating the gravity equation using data on worker flows across occupations only for individuals aged 18 to 35. Results obtained when using this sample of younger workers are quite consistent with the findings from the full sample. Columns (1) and (2) of Table 10 present results of the counterfactual experiments analogous to those in Table 8. Although there is more dispersion in the estimated role of task content across occupations, the effects for the median occupation are remarkably similar to those estimated from the full sample.

We also estimate our gravity equation using flows for workers with high levels of education (college graduates), as many of these workers may have the option to match with a wide variety of occupations. The results are presented in Columns (3) and (4) of Table 10, and suggest an even weaker role of task content for this group of workers.[39]

5.4 Occupation Coding Error

Over our sample period the CPS consistently used dependent coding techniques to assign occupation codes.[40] Moscarini and Thomsson (2007) discuss how dependent coding reduces the number of spurious transitions due to mis-coding. However, as emphasized by Kambourov and Manovskii (2013), the dependent coding procedure is only applied to workers who remain with the same employer. Therefore, mis-coding may still occur for workers who switch

[39]If we focus instead on workers without a college degree, we find that task-related costs account for 7% of total costs for the median occupation, and 8% on average across all occupation-year cells.

[40]Dependent coding involves importing information collected in the previous month's interview into the current interview. Instead of asking individuals to verbally describe their current occupation in each month and independently coding these descriptions, interviewees were asked whether they still had the same job as in the previous month. If so, they were automatically assigned the same occupation code. The introduction of dependent coding in 1994 has a dramatic effect on measured occupation mobility rates, as illustrated in Appendix Figure A.4.

employers in the post-1994 period.

In order to obtain specific evidence on the effect of coding error, we contrast the estimates of our model for 1994 (the first year in which dependent coding techniques were used in the CPS) to those obtained using data for 1992, prior to the introduction of these techniques.[41] The comparison allows us to determine the likely effect of coding error on the estimated coefficients of the task variables. Results from the counterfactual experiments are shown in Table 11. In 1994, after the introduction of dependent coding techniques, we find that task-related costs account for 14% of total transition costs on average across occupations. The corresponding figure for the year 1992, before dependent coding techniques were introduced, is substantially higher at 38%.

This evidence implies that coding error tends to lead to an overestimation of the relative importance of task content heterogeneity. This finding can be understood by considering the nature of occupation mis-classification due to coding error. Mis-classification occurs when an individual's description of the same job in two consecutive periods is assigned a different occupation code. In such cases, the erroneous code will likely correspond to an occupation that is fairly similar – in terms of task content – to the original occupation. Hence coding error will predominantly lead to spurious transitions across occupation pairs with similar task content.[42] This implies that in the presence of coding error, our approach would over-estimate the relative importance of task content, and therefore, to the extent that some coding error remains in the post-1994 period, the relative importance of the task variables may be even smaller than what we find.

6 Discussion

Quantifying the costs faced by workers when considering an occupational transition is key to understand the aggregate impact of labor market shocks and their distributional consequences. Technological change or trade reforms may reshuffle labor demand across occupations and trigger adjustments which critically depend on the magnitude of mobility costs. Arguably, large mobility costs would impose disproportionate losses on workers whose current occupation suffers from decreased demand.

Our novel approach allows us to gauge the magnitude of transition costs across any pair of occupations. A crucial advantage of this approach is that identification relies primarily on worker flows, rather than wage data, allowing for a broad notion of payoffs and costs (whether

[41] Matching of individuals across months is problematic during the last months of 1993; hence, we use 1992 data for this exercise in order to have flow data for a full calendar year.

[42] Data from the Panel Study of Income Dynamics (PSID) show a striking discontinuity in median distance between the period where occupations were retrospectively coded, and a more recent period where the incidence of coding error is higher; see Figure 3.1 in Cortes (2012). Robinson (2011) also provides a discussion of these issues.

pecuniary or non-pecuniary). Through counterfactual experiments we show that transition costs across most occupation pairs are very large: if all pair-wise costs were reduced to the lowest value estimated in our sample, the aggregate mobility rate across 2-digit occupations would increase by an order of magnitude. The large transition costs imply that a negative demand shock that persistently reduces payoffs in a particular occupation will not induce much additional worker reallocation, as many workers will find it optimal to remain in that occupation and avoid incurring these costs. This is consistent with evidence in Autor and Dorn (2009) who find that routine jobs – which experience a decline in demand due to new automation technologies – are "getting old".[43] Our findings also complement influential work on labor market mobility which, using different identification strategies, has highlighted the large magnitude of the costs faced by workers considering a transition across industries (see Dix-Carneiro, 2014; Artuc et al., 2010).

Existing studies on industry-switching costs typically draw inference from data on sectoral wages; hence, the identification of intersectoral mobility costs critically depends on the quality of counterfactual wage differentials measured across sectors, and their expected evolution. As noted by Dix-Carneiro (2014), this line of research only offers limited insights into the nature of the costs faced by workers when switching sectors. Yet, an understanding of what constitutes mobility costs is necessary to make sense of observed patterns and to inform policy. To that end, a key contribution of our paper is to explore the relative importance of task heterogeneity as a component of occupational mobility costs. Our estimates suggest that task content heterogeneity has a significant impact on transition costs: a one standard deviation increase in task distance raises the cost of switching jobs by approximately 14%. However, we also find that most mobility costs are not related to task differences and can be best described as occupation-specific access costs. These access costs account for at least 80% of overall transition costs between most occupation pairs. Thus, they play a critical role in the labor adjustment process and, as we show, can be partly related to existing labor market institutions. Incidentally, our findings corroborate those of several studies looking at workers' reallocation patterns following trade liberalizations. In this context, Ritter (2014) shows that labor market institutions play a larger role in the adjustment process than specific human capital, confirming earlier results in Kambourov (2009). Similarly, Dix-Carneiro (2014) argues that mobility costs other than sector-specific experience are of crucial importance, and may explain the slow adjustment of the labor market following trade reforms in Brazil.

The literature on task human capital (e.g. Poletaev and Robinson, 2008; Gathmann and Schönberg, 2010) has identified task tenure as a significant component of individual wage

[43] Evidence in Cortes et al. (2014) also highlights the importance of inflows from unemployment and non-participation when accounting for the decline in routine employment, suggesting that the economy has adjusted to the decreased demand for routine jobs through reduced inflows rather than increased outflows. Our findings indicate that this asymmetry can be rationalized by the high cost of transiting out of routine occupations for incumbents.

growth. This is not inconsistent with our results. However, we find that the cost of switching occupations considerably exceeds the loss of task-specific human capital. Hence, researchers working on models motivated by the task human capital literature should be cautious about assuming that transition costs depend exclusively on the extent of overlap in task content between occupations. Calibrating these costs based only on task distance measures would be misleading. Realistic portraits of the labor market should carefully account for formal or implicit institutional features that constrain access to occupations, severely limiting labor movement even across occupations with substantial overlap in terms of their task content. These considerations are especially relevant for policy analysis.

It is important to acknowledge that certain caveats apply to our results. The empirical analysis is based on transitions occurring over one-month horizons; hence our estimates should be interpreted as describing occupational transition costs over relatively short intervals.[44] Moreover, we emphasize the transferability of task-specific human capital and, therefore, we focus on the overlap in task content within each occupation pair. A broader notion of task human capital could encompass all the skills that can be attributed to task content, regardless of their portability to other occupations. For example, particular bundles of tasks could affect payoffs in different occupations by impacting the mean level of productivity of incumbent workers, or the speed of human capital accumulation in those occupations. Through these channels, task content could influence occupational mobility in ways not directly related to skill transferability as captured by task distance.[45] To explore these complementary mechanisms one would have to specify a richer, but rather more restrictive, life-cycle model of productivity and wages.

7 Conclusions

We quantify the costs of occupational mobility using an approach which relies on data on worker flows across occupations rather than wages. This approach circumvents potentially confounding effects embedded in wage changes observed for occupational switchers, and allows us to estimate the total costs (both pecuniary and non-pecuniary) that workers would face if they chose to change occupations. These are the costs that limit mobility, and they are not equal to the costs that are actually incurred by workers who decide to make a switch.

We posit an occupational choice model where workers draw a match quality shock for each potential occupation from a set of extreme value distributions, and choose optimally which occupation to work in, based on their draws, the utility payoffs of alternative occupations and

[44] Unfortunately our data is not suitable for an analysis of longer-term transitions due to the presence of coding error, as discussed in Appendix G.

[45] It is important to stress that our framework is not incompatible with the possibility that certain bundles of tasks are associated with a faster rate of human capital accumulation. If this accumulated human capital is task-specific, it should remain portable across occupations with similar task content.

the costs of moving out of their current occupation. The model naturally maps into a gravity specification linking worker flows to occupation characteristics, and to the implicit transition costs faced by workers.

Our empirical analysis quantifies different layers of occupational transition costs. In particular, we assess the role of task distance (the degree of dissimilarity in the mix of task requirements) as a component of the cost of switching among any two occupations. We find that raising task distance by one standard deviation increases the cost of switching occupations by approximately 14% in our baseline specification, all else equal. In addition, if the switch involves moving across major task groups mobility costs are raised much further, in ways not captured by the pure distance measure. Yet, despite the considerable role of task content, we find that the largest share of occupational mobility costs is attributable to task-independent factors, modeled as occupation entry costs in our baseline specification. These occupation-specific costs vary widely in size, and are positively correlated with measures of training, unionization and licensing requirements at the occupational level. Overall, estimated transition costs across occupations are substantial and remain fairly stable over the period that we consider. The results are robust to alternative ways of characterizing the task content of occupations, using both Dictionary of Occupational Titles and O*Net data, and hold also when we focus exclusively on the mobility patterns of younger workers, or workers with a college degree.

Our findings complement the evidence regarding transition costs across industries in Artuc et al. (2010), Artuc and McLaren (2015) and Dix-Carneiro (2014), and caution against assuming perfect mobility of workers across occupations, or positing that the cost of switching is homogeneous and/or mainly driven by loss of task-specific human capital. Models that make such assumptions abstract from transition costs that are very sizable and heterogeneous across occupations.

Future work might embed our flow model within an equilibrium framework featuring shocks to demand for different occupations (due to technology or trade, as in Tombe and Zhu, 2015). The occupational sorting mechanism in our model would provide a parsimonious setup to study worker flows across occupations whose fortunes change over time. With finer data, the model could also encompass richer heterogeneity in terms of match quality distributions or mobility costs. It would also be interesting to consider a richer dynamic setting where workers explicitly maximize a present discounted value of lifetime utility, and some assumption is made about the persistence of match-quality draws. As stressed by Head and Mayer (2013) in their review article, this is an extension that micro-founded gravity models have yet to address and remains a promising direction for future work.

References

Acemoglu, Daron and David Autor (2011), "Skills, tasks and technologies: Implications for employment and earnings." *Handbook of Labor Economics*, 4, 1043–1171.

Artuc, Erhan, Shubham Chaudhuri, and John McLaren (2010), "Trade shocks and labor adjustment: A structural empirical approach." *American Economic Review*, 100, 1008–1045.

Artuc, Erhan and John McLaren (2015), "Trade policy and wage inequality: A structural analysis with occupational and sectoral mobility." *Journal of International Economics*, 97, 278 – 294.

Autor, David and David Dorn (2009), "This job is "getting old": Measuring changes in job opportunities using occupational age structure." *American Economic Review*, 99, 45–51.

Autor, David H. and David Dorn (2013), "The growth of low skill service jobs and the polarization of the U.S. labor market." *American Economic Review*, 103, 1553–1597.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2006), "The polarization of the US labor market." *The American Economic Review*, 96, 189–194.

Autor, David H., Frank Levy, and Richard J. Murnane (2003), "The skill content of recent technological change: An empirical exploration." *Quarterly Journal of Economics*, 118, 1279–1333.

Caines, Colin, Florian Hoffmann, and Gueorgui Kambourov (2015), "Training programs, skills, and human capital: A life-cycle approach." *Working Paper.*

Cortes, G. M., N. Jaimovich, C. J. Nekarda, and H. E. Siu (2014), "The Micro and Macro of Disappearing Routine Jobs: A Flows Approach." *NBER Working Paper No. 20307.*

Cortes, Guido Matias (2012), "Essays on the task content of occupations and occupational mobility." *University of British Columbia Dissertation.*

Cortes, Guido Matias (2016), "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data." *Journal of Labor Economics*, 34, 63–105.

Dix-Carneiro, Rafael (2014), "Trade liberalization and labor market dynamics." *Econometrica*, 82, 825–885.

Eaton, J. and S. Kortum (2002), "Technology, geography, and trade." *Econometrica*, 70, 1741–1779.

Eaton, Jonathan and Samuel Kortum (2001), "Trade in capital goods." *European Economic Review*, 45, 1195–1235.

Flood, Sarah, Miriam King, Steven Ruggles, and J. Robert Warren (2015), "Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-readable database]." *Minneapolis: University of Minnesota.*

Gathmann, Christina and Uta Schönberg (2010), "How general is human capital? A task-based approach." *Journal of Labor Economics*, 28, 1–49.

Gervais, Martin, Nir Jaimovich, Henry E Siu, and Yaniv Yedid-Levi (2016), "What should I be when I grow up? Occupations and unemployment over the life cycle." *Forthcoming, Journal of Monetary Economics.*

Gibbons, Robert, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent (2005), "Comparative advantage, learning, and sectoral wage determination." *Journal of Labor Economics*, 23, 681–724.

Goos, Maarten and Alan Manning (2007), "Lousy and lovely jobs: The rising polarization of work in Britain." *The Review of Economics and Statistics*, 89, 118–133.

Gottschalk, Peter, David A Green, and Benjamin M Sand (2015), "Taking selection to task: Bounds on trends in occupational task prices for the US, 1984-2013." *Working Paper.*

Groes, Fane, Philipp Kircher, and Iourii Manovskii (2015), "The U-shapes of occupational mobility." *The Review of Economic Studies*, 82, 659–692.

Head, Keith and Thierry Mayer (2013), "Gravity equations: Workhorse, toolkit, and cookbook." *Handbook of International Economics*, 4, 131–195.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding (2010), "Inequality and unemployment in a global economy." *Econometrica*, 78, 1239–1283.

Herz, Benedikt and Thijs Van Rens (2015), "Accounting for mismatch unemployment." *Working Paper.*

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow (2013), "The Allocation of Talent and US Economic Growth." *NBER Working Paper No. 18693.*

ICPSR, 1981 (1981), *Dictionary of Occupational Titles (DOT): Part I - Current Population Survey, April 1971, Augmented with DOT Characteristics, and Part II - Fourth Edition Dictionary of DOT Scores for 1970 Census Categories [Computer File]. Study Nr. 07845.* National Academy of Sciences - Committee on Occupational Classification and Analysis, Washington DC: U.S. Department of Commerce, Bureau of the Census [Producer], Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor].

ICPSR, 1991 (1991), *Dictionary of Occupational Titles (DOT): Revised Fourth Edition, 1991 [Computer file]. Study Nr. 06100. United States Department of Labor, United States Employment Service, and the North Carolina Occupational Analysis Field Center*. United States Department of Labor, United States Employment Service, and the North Carolina Occupational Analysis Field Center, Washington, DC: United States Department of Labor, United States Employment Service, and Raleigh, NC: North Carolina Occupational Analysis Field Center [Producer], 1991. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor], 1994.

Ingram, B.F. and G.R. Neumann (2006), "The returns to skill." *Labour Economics*, 13, 35–59.

Kambourov, G. and I. Manovskii (2013), "A Cautionary Note on Using (March) CPS and PSID Data to Study Worker Mobility." *Macroeconomic Dynamics*, 17, 172–194.

Kambourov, Gueorgui (2009), "Labour Market Regulations and the Sectoral Reallocation of Workers: The Case of Trade Reforms." *Review of Economic Studies*, 76, 1321–1358.

Kambourov, Gueorgui and Iourii Manovskii (2008), "Rising occupational and industry mobility in the United States: 1968-1997." *International Economic Review*, 49, 41–79.

Kambourov, Gueorgui and Iourii Manovskii (2009a), "Occupational mobility and wage inequality." *Review of Economic Studies*, 76, 731–759.

Kambourov, Gueorgui and Iourii Manovskii (2009b), "Occupational specificity of human capital." *International Economic Review*, 50, 63–115.

Lalé, Etienne (2015), "Worker reallocation across occupations: Confronting data with theory." *Working Paper*.

Lazear, Edward P (2009), "Firm-specific human capital: A skill-weights approach." *Journal of Political Economy*, 117, 914–940.

Lemieux, Thomas (2006), "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?" *The American Economic Review*, 461–498.

Meyer, PB and AM Osborne (2005), "Proposed Category System for 1960-2000 Census Occupations." *Bureau of Labor Statistics Working Paper*, 383.

Moscarini, Giuseppe and Kaj Thomsson (2007), "Occupational and job mobility in the US." *Scandinavian Journal of Economics*, 109, 807–836.

Neal, D. (1995), "Industry-specific human capital: Evidence from displaced workers." *Journal of Labor Economics*, 13, 653–677.

Neal, Derek A (1999), "The complexity of job mobility among young men." *Journal of Labor Economics*, 17, 237–61.

Nekarda, Christopher J (2009), "A Longitudinal Analysis of the Current Population Survey: Assessing the Cyclical Bias of Geographic Mobility." *Federal Reserve Board of Governors*.

Parent, D. (2000), "Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics." *Journal of Labor Economics*, 18, 306–323.

Pavan, Ronni (2011), "Career choice and wage growth." *Journal of Labor Economics*, 29, 549–587.

Peri, G. and C. Sparber (2009), "Task specialization, immigration, and wages." *American Economic Journal: Applied Economics*, 1, 135–169.

Poletaev, Maxim and Chris Robinson (2008), "Human capital specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000." *Journal of Labor Economics*, 26, 387–420.

Ritter, Moritz (2014), "Offshoring and occupational specificity of human capital." *Review of Economic Dynamics*, 17, 780–798.

Robinson, Chris (2011), "Occupational mobility, occupation distance and specific human capital." *University of Western Ontario, CIBC Centre for Human Capital and Productivity Working Papers*.

Schoellman, Todd (2012), "Education quality and development accounting." *The Review of Economic Studies*, 79, 388–417.

Shaw, Kathryn L (1984), "A formulation of the earnings function using the concept of occupational investment." *Journal of Human Resources*, 319–340.

Sorkin, Isaac (2015), "Ranking firms using revealed preference." *Working Paper*.

Sullivan, Paul (2010), "Empirical evidence on occupation and industry specific human capital." *Labour Economics*, 17, 567–580.

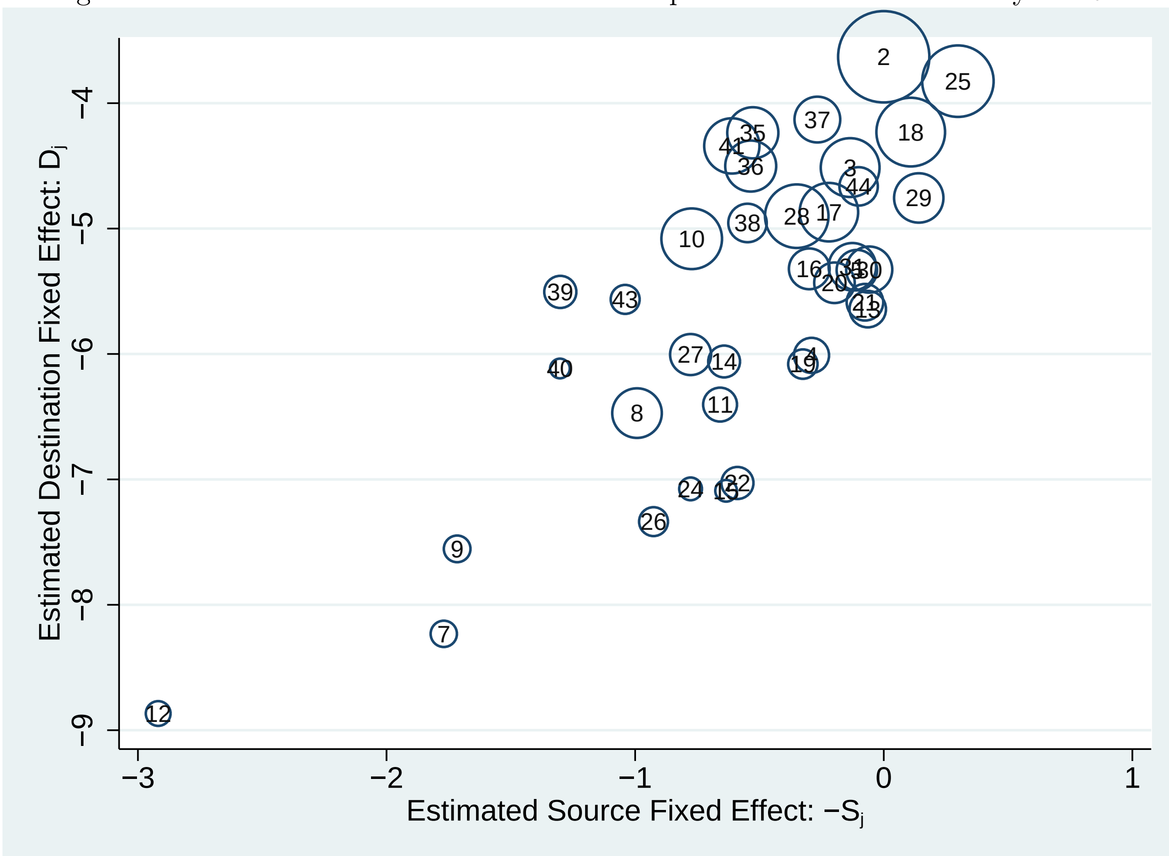
Tombe, Trevor and Xiaodong Zhu (2015), "Trade, Migration and Regional Income Differences: Evidence from China." 2015 meeting papers, Society for Economic Dynamics.

Topel, R. (1991), "Specific capital, mobility, and wages: Wages rise with job seniority." *Journal of Political Economy*, 99, 145.

Willis, Robert J (1986), "Wage determinants: A survey and reinterpretation of human capital earnings functions." *Handbook of Labor Economics*, 1, 525–602.

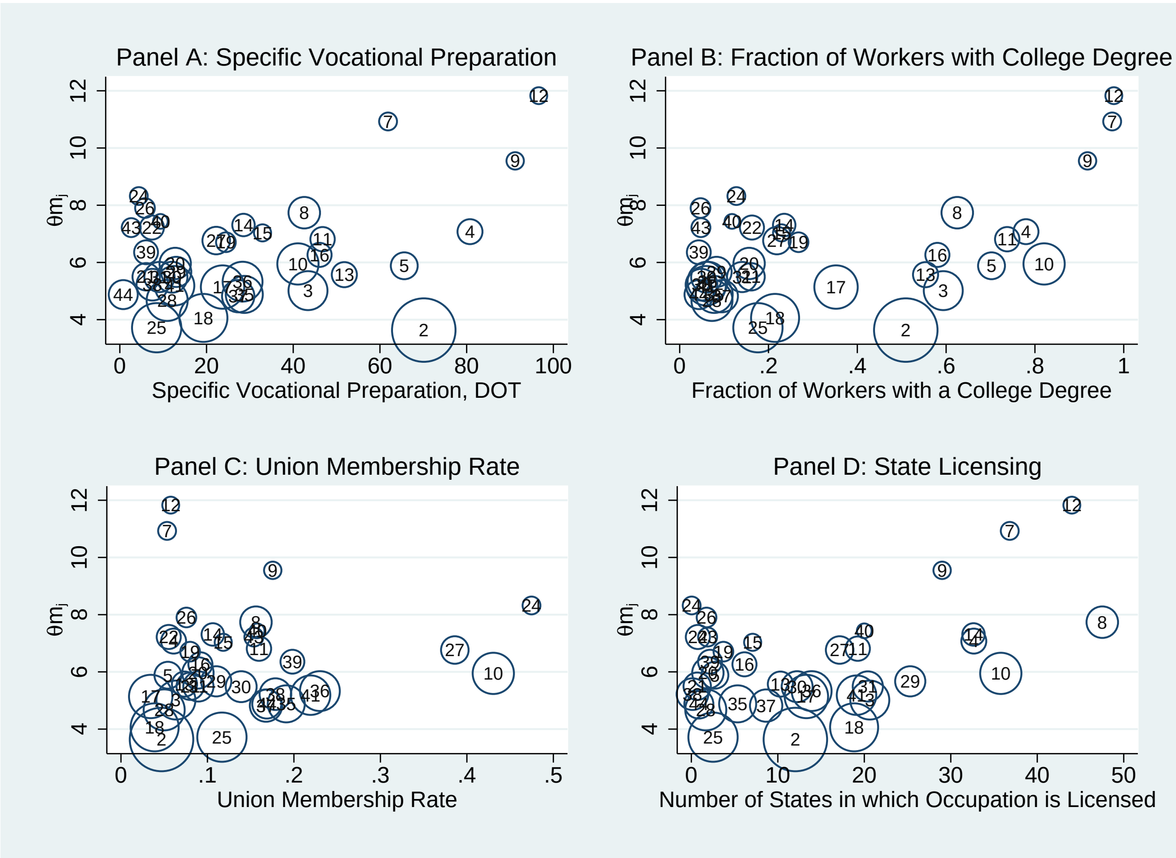
Yamaguchi, Shintaro (2012), "Tasks and heterogeneous human capital." *Journal of Labor Economics*, 30, pp. 1–53.

Figure 1: Estimated source and destination occupation fixed effects for the year 2012



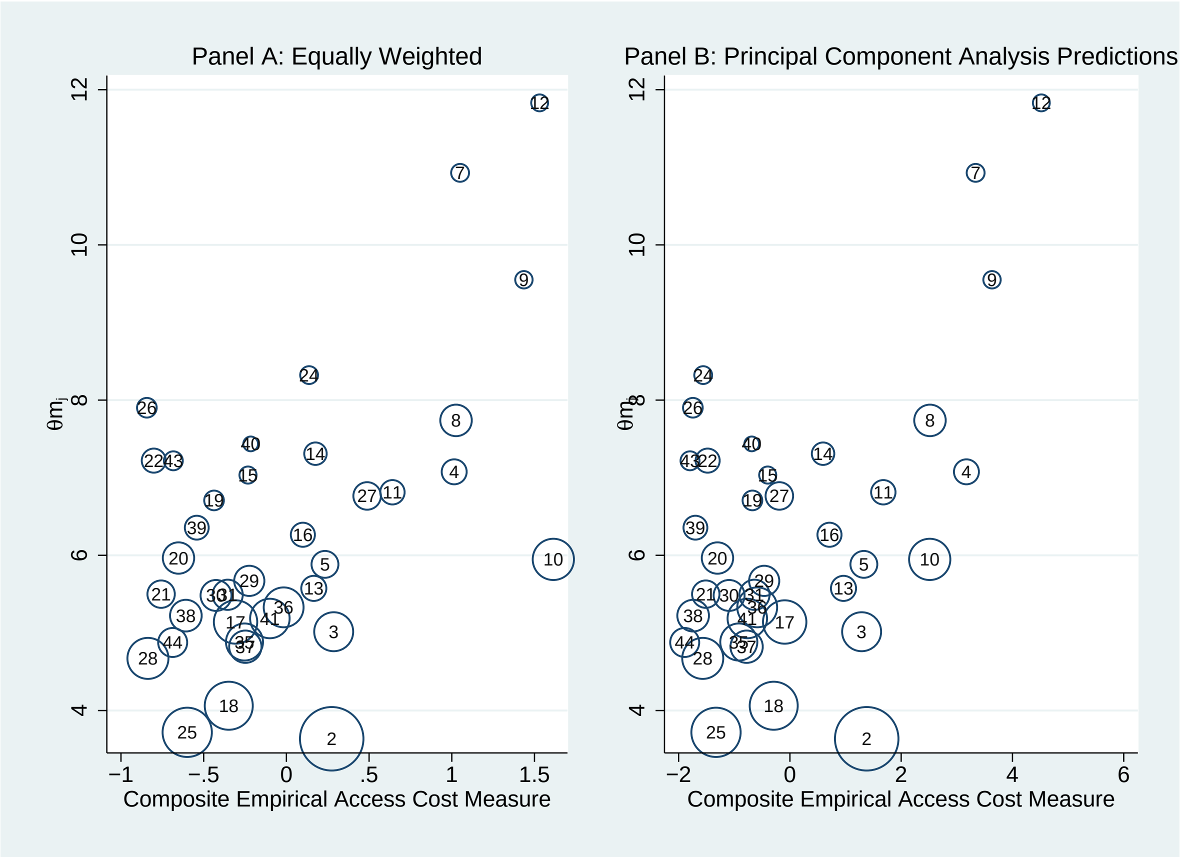
Note: Each circle represents a 2-digit occupation, with the size of the circle representing the size of the occupation in 2012, and labeled with its corresponding occupation code; see Appendix Table A.1 for the definitions of each occupation code.

Figure 2: Correlation between estimates of θm_j and other empirical measures of occupation access costs



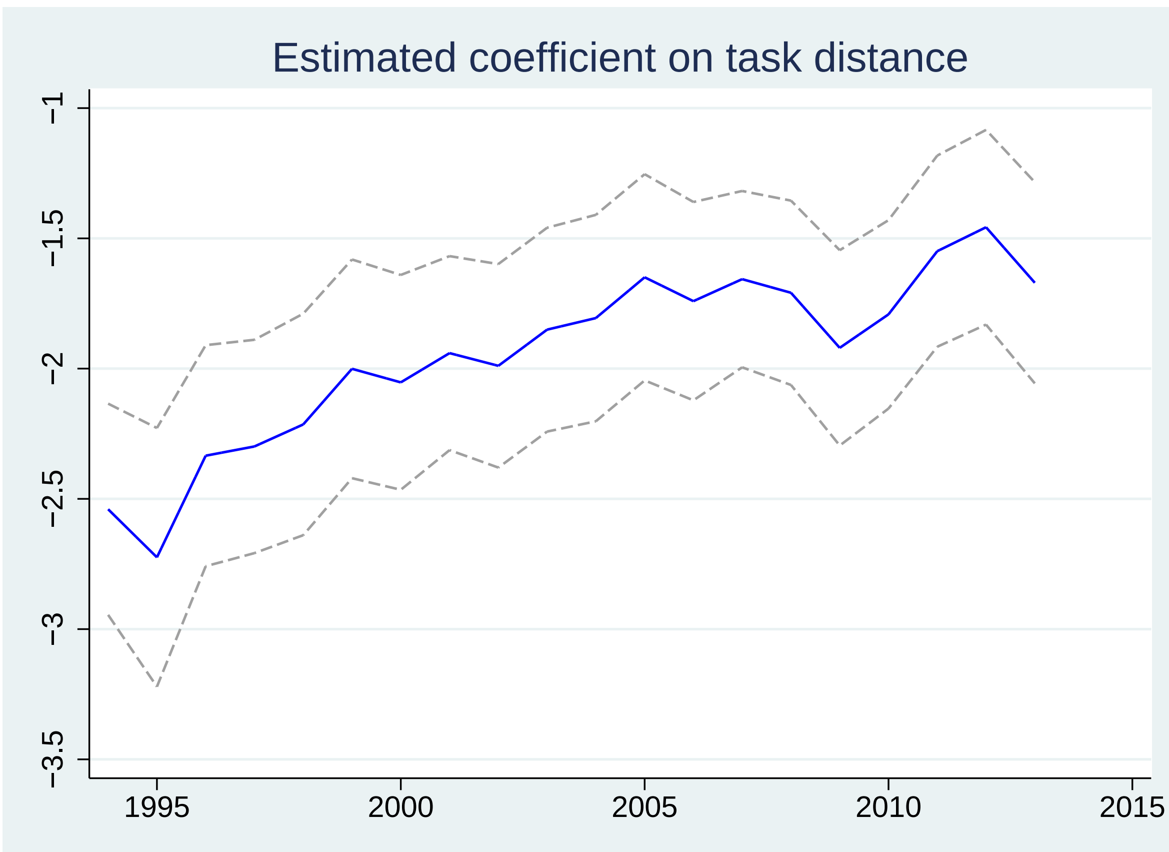
Note: Each circle represents a 2-digit occupation, with the size of the circle representing the average size of the occupation across all years in our sample, and labeled with its corresponding occupation code; see Appendix Table A.1 for the definitions of each occupation code.

Figure 3: Correlation between estimates of θm_j and composite empirical measures of occupation access costs



Note: Each circle represents a 2-digit occupation, with the size of the circle representing the average size of the occupation across all years in our sample, and labeled with its corresponding occupation code; see Appendix Table A.1 for the definitions of each occupation code.

Figure 4: Evolution of the estimated coefficient on task distance over time



Note: The dashed lines indicate a 95% confidence interval.

Table 1: Examples of Task Content

	Health Assessment and Treating	Cleaning and Building Service	Helpers, Construction and Production	Freight, Stock and Material Handlers
GED-Reasoning	1.12	-0.79	-1.52	-1.92
GED-Math	1.08	-0.59	-1.24	-1.42
GED-Language	1.18	-0.61	-1.33	-1.59
Intelligence	0.91	-0.63	-0.87	-0.90
Verbal	0.83	-0.60	-0.85	-0.84
Numerical	0.57	-0.57	-0.82	-0.91
Spacial	0.42	-0.39	-0.51	-0.73
Form Perception	1.07	-0.63	-0.68	-0.83
Clerical	0.88	-0.58	-0.98	-1.01
Motor Coord	1.05	-0.64	-0.46	-0.72
Finger Dext	0.66	-0.57	-0.50	-0.59
Manual Dext	0.68	-0.17	0.06	0.00
Eye-Hand-Foot	-0.05	0.23	-0.04	-0.30
Color Discrim	0.56	-0.34	-0.54	-0.73
Distance (DOT)	0.986		0.004	
Distance (ONet)	0.807		0.062	

Note: Each DOT dimension is normalized to have mean zero and standard deviation one across the universe of standardized 3-digit occupations from Autor and Dorn (2013). More details are provided in Appendix C. Distance is calculated as in Equation (13). The distance based on O*Net uses the work activities listed in Appendix Table A.3.

Table 2: Estimated coefficients on 'gravity-type' equation, 2012

	No Zeros	Zeros Replaced	
	OLS	OLS	IntReg
	(1)	(2)	(3)
dist	-1.132	-1.394	-1.457
	(.107)***	(.180)***	(.191)***
λ^{NC}	-.182	-.628	-.690
	(.120)	(.204)***	(.216)***
λ^{RC}	-.584	-.686	-.676
	(.137)***	(.236)***	(.249)***
λ^{RM}	-1.209	-1.629	-1.672
	(.132)***	(.226)***	(.239)***
λ^{NM}	-.784	-.730	-.712
	(.203)***	(.354)**	(.374)*
Const.	-4.467	-3.737	-3.630
	(.193)***	(.338)***	(.356)***
Obs.	1218	1332	1332
R^2	.588	.515	

Note: The table presents the results from the estimation of Equation (15) for the year 2012. Observations are at the occupation pair level. The dependent variable is the normalized flow of workers $\ln(sw_{kj}/sw_{kk})$. All specifications include source and destination occupation dummies.

Table 3: Summary statistics of estimates of θ

Sample:	Full	Full	Young
Measure:	Wages	Residual Wages	Wages
	(1)	(2)	(3)
10th Percentile	2.182	2.504	2.437
25th Percentile	2.435	2.725	2.769
50th Percentile	2.851	3.023	3.230
75th Percentile	3.151	3.304	3.770
90th Percentile	3.411	3.530	4.350
Mean	2.820	3.019	3.333
Nr of Cells	7,448	7,441	8,505

Note: The table presents summary statistics for the estimated value of θ based on the standard deviation of wages or residual wages as described in Section 4.3. The estimation in Columns (1) and (2) uses the full sample in each month, by initial occupation, excluding occupation-month cells with less than 100 observations; the estimation in Column (3) uses a restricted sample of gender-specific demographic bins for those aged 25 to 30, by month and by initial occupation, excluding cells with less than 15 observations.

Table 4: Estimated effects on occupational transition costs for each of the specifications in Table 2

	Implied $\widehat{\beta}$			Percentage effect on d_{kj}		
	No Zeros	Zeros Replaced		No Zeros	Zeros Replaced	
	OLS	OLS	IntReg	OLS	OLS	IntReg
	(1)	(2)	(3)	(4)	(5)	(6)
$dist$	0.351	0.432	0.451	10.643	13.441	14.086
λ^{NC}	0.056	0.194	0.214	5.790	21.448	23.833
λ^{RC}	0.181	0.212	0.209	19.828	23.647	23.265
λ^{RM}	0.374	0.504	0.518	45.393	65.612	67.784
λ^{NM}	0.243	0.226	0.221	27.459	25.351	24.675

Note: Columns (1) to (3) show the marginal effects of distance and the task group switching dummies on the normalized flow of workers across occupation pairs using the estimate of $\theta = 3.23$. Columns (4) to (6) show the percentage effect on the transition cost d_{kj} from a one standard deviation increase in distance and from a change from 0 to 1 for the task group switching variables. The standard deviation of distance is computed among the sample of occupation pairs with non-zero flows for the purposes of Column (4) and among the full sample of occupation pairs for the purposes of Columns (5) and (6).

Table 5: Estimated occupational transition costs between selected occupation pairs, 2012

Source	Destination	Estimated d_{kj}		
Occupation	Occupation	Distance	Tasks	All costs
Info and records processing, excl financial	Other admin support occ, incl clerical	1.005	1.005	2.994
Office supervisors and computer operators	Other admin support occ, incl clerical	1.031	1.031	3.071
Management related occupations	Executives, administrators and managers	1.005	1.005	3.091
Librarians, social scientists, religious workers	Executives, administrators and managers	1.006	1.006	3.095
Financial records processing occupations	Other admin support occ, incl clerical	1.043	1.043	3.107
Teachers, except college and university	Executives, administrators and managers	1.015	1.015	3.122
Teachers, college and university	Executives, administrators and managers	1.016	1.016	3.127
Lawyers and judges	Executives, administrators and managers	1.023	1.023	3.146
Office machine operators and mail distributing	Other admin support occ, incl clerical	1.060	1.060	3.158
Retail and other salespersons	Other admin support occ, incl clerical	1.065	1.065	3.174
Food service occupations	Lawyers and judges	1.416	1.754	67.389
Helpers, construction and production occ	Lawyers and judges	1.432	1.774	68.160
Transportation and material moving	Lawyers and judges	1.434	1.775	68.219
Other personal service occupations	Lawyers and judges	1.434	1.776	68.250
Production inspectors and graders	Lawyers and judges	1.446	1.791	68.813
Mechanics and repairers	Lawyers and judges	1.461	1.810	69.541
Machine operators and tenders, not precision	Lawyers and judges	1.462	1.810	69.567
Fabricators, assemblers and hand working occ	Lawyers and judges	1.477	1.829	70.298
Construction trades	Lawyers and judges	1.480	1.832	70.410
Other precision production occupations	Lawyers and judges	1.508	1.868	71.766

Table 6: Summary statistics for the relative size of the transition cost associated with the task-related variables

	Distance	Tasks
	(1)	(2)
10th Percentile	0.005	0.010
25th Percentile	0.015	0.027
50th Percentile	0.031	0.057
75th Percentile	0.050	0.095
90th Percentile	0.069	0.131
Maximum	0.229	0.283
Mean	0.036	0.065
Obs.	26,640	26,640

Note: The observations are occupation pair-year cells. Column (1) presents the summary statistics for the fraction of the transition costs that can be attributed to task distance, while Column (2) presents the fraction that can be attributed to all task-related barriers (task distance and costs of transitioning across broad task groups). The remainder is accounted for by task-independent occupational entry costs.

Table 7: Observed and counterfactual occupational mobility rates for selected occupations, 2012

	Mobility						
Occupation	Observed	Fitted	No dist	No task	Min costs	Union	Licensing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lawyers and judges	0.023	0.004	0.007	0.012	0.156	0.005	0.006
Health assessment and treating occupations	0.028	0.023	0.044	0.079	0.610	0.027	0.033
Health diagnosing occupations	0.029	0.010	0.021	0.038	0.408	0.011	0.014
Teachers, except college and university	0.029	0.036	0.052	0.095	0.662	0.040	0.050
Protective service occupations	0.033	0.030	0.054	0.096	0.661	0.035	0.044
Teachers, college and university	0.036	0.015	0.022	0.040	0.422	0.017	0.022
Transportation and material moving	0.039	0.044	0.073	0.107	0.698	0.059	0.059
Other personal service occupations	0.040	0.048	0.081	0.167	0.791	0.058	0.069
Librarians, social scientists, religious workers	0.041	0.042	0.061	0.107	0.688	0.047	0.060
Food service occupations	0.042	0.038	0.064	0.137	0.751	0.047	0.059
Construction trades	0.057	0.048	0.079	0.115	0.714	0.064	0.063
Retail and other salespersons	0.057	0.069	0.104	0.196	0.827	0.084	0.092
Other administrative support occupations, inc clerical	0.060	0.076	0.116	0.222	0.853	0.084	0.115
Fabricators, assemblers and hand working occ	0.060	0.026	0.040	0.059	0.531	0.035	0.036
Machine operators and tenders, not precision	0.061	0.052	0.080	0.116	0.711	0.070	0.073
Engineering and science technicians	0.065	0.031	0.062	0.110	0.693	0.036	0.044
Office supervisors and computer operators	0.070	0.053	0.078	0.143	0.755	0.063	0.074
Information and records processing, except financial	0.071	0.068	0.097	0.175	0.799	0.082	0.095
Freight, stock and material handlers	0.083	0.077	0.118	0.168	0.795	0.101	0.107
Helpers, construction and production occ	0.099	0.034	0.052	0.075	0.598	0.045	0.047
Aggregate	0.049	0.050	0.081	0.145	0.744	0.060	0.072

Note: In Column (3) counterfactual occupational mobility rates are calculated for the case when task distance is reduced to zero. In Column (4), transition costs across broad task groups are also reduced to zero. In Column (5), task-independent occupational entry costs are reduced to their lowest estimated value in the sample. In Column (6), access costs are reduced for occupations with above-median membership rates. In Column (7), access costs are reduced for occupations with above-median licensing rates.

Table 8: Summary statistics for the fraction of the transition costs that can be attributed to task-related variables based on the results from the counterfactual experiments

	Distance Fraction	Tasks Fraction
	(1)	(2)
10th Percentile	0.026	0.060
25th Percentile	0.034	0.081
50th Percentile	0.045	0.105
75th Percentile	0.057	0.134
90th Percentile	0.068	0.163
Maximum	0.133	0.267
Mean	0.047	0.110
Obs.	740	740

Note: The observations are occupation-year cells. Column (1) presents the summary statistics for the fraction of the counterfactual increase in mobility that can be attributed to task distance, while Column (2) presents the fraction that can be attributed to all task-related barriers (task distance and costs of transitioning across broad task groups). The remainder is accounted for by heterogeneity in task-independent occupational entry costs.

Table 9: Robustness checks for the results on the relative size of the transition cost associated with the task-related variables

	$\theta = 2$		$\theta = 2.437$		$\theta = 4.350$		$\theta = 8.87$	
	Distance	Tasks	Distance	Tasks	Distance	Tasks	Distance	Tasks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
10th Percentile	.002	.003	.003	.005	.008	.016	.014	.029
25th Percentile	.005	.010	.008	.016	.023	.040	.042	.071
50th Percentile	.011	.022	.018	.035	.046	.081	.077	.132
75th Percentile	.020	.044	.031	.064	.071	.128	.114	.193
90th Percentile	.031	.064	.046	.090	.094	.172	.146	.250
Maximum	.140	.182	.178	.226	.275	.333	.352	.415
Mean	.015	.029	.023	.043	.050	.089	.080	.136
Obs.	26,640	26,640	26,640	26,640	26,640	26,640	26,640	26,640

Note: The observations are occupation pair-year cells. Columns (1), (3), (5) and (7) present the summary statistics for the fraction of the transition costs that can be attributed to task distance using the value of θ indicated on the first row. Columns (2), (4), (6) and (8) present the fraction that can be attributed to all task-related barriers (task distance and costs of transitioning across broad task groups). The remainder is accounted for by heterogeneity in task-independent occupational entry costs.

Table 10: Results from the counterfactual experiments using alternative sub-samples

Sample:	Young		College	
	Distance Fraction	Tasks Fraction	Distance Fraction	Tasks Fraction
	(1)	(2)	(3)	(4)
10th Percentile	.018	.043	.002	.004
25th Percentile	.031	.074	.005	.013
50th Percentile	.051	.122	.011	.034
75th Percentile	.073	.169	.023	.078
90th Percentile	.096	.224	.042	.149
Maximum	.177	.379	.294	.621
Mean	.055	.128	.018	.058
Obs.	740	740	740	740

Note: The observations are occupation-year cells. The results are based on the estimation of Equation (15) using data for younger workers only (aged 18 to 35) in Columns (1) and (2), and using data for college graduates only in Columns (3) and (4). Columns (1) and (3) present the summary statistics for the fraction of the counterfactual increase in mobility that can be attributed to task distance, while Columns (2) and (4) present the fraction that can be attributed to all task-related barriers (task distance and costs of transitioning across broad task groups). The remainder is accounted for by heterogeneity in task-independent occupational entry costs.

Table 11: Results from the counterfactual experiments before and after the introduction of dependent coding techniques

	Dependent Coding (1994)		No Dependent Coding (1992)	
	Distance Fraction	Tasks Fraction	Distance Fraction	Tasks Fraction
	(1)	(2)	(3)	(4)
10th Percentile	.037	.078	.077	.254
25th Percentile	.049	.109	.129	.287
50th Percentile	.063	.141	.158	.380
75th Percentile	.070	.167	.214	.492
90th Percentile	.099	.190	.251	.602
Maximum	.124	.275	.374	.692
Mean	.064	.141	.171	.402
Obs.	37	37	37	37

Note: The observations are at the occupation level. The results are based on the estimation of Equation (15) using data for 1994 (after the introduction of dependent coding techniques) in Columns (1) and (2), and using data for 1992 (before the introduction of dependent coding techniques) in Columns (3) and (4). Columns (1) and (3) present the summary statistics for the fraction of the counterfactual increase in mobility that can be attributed to task distance, while Columns (2) and (4) present the fraction that can be attributed to all task-related barriers (task distance and costs of transitioning across broad task groups). The remainder is accounted for by heterogeneity in task-independent occupational entry costs.

Online Appendix for:

"The Costs of Occupational Mobility: An Aggregate Analysis"

Guido Matias Cortes (University of Manchester and RCEA)
Giovanni Gallipoli (University of British Columbia, HCEO and RCEA)

Appendix A Extension: Occupation Tenure

This section extends the model to allow for occupation-specific human capital.[1] Let an individual's tenure in occupation j be denoted $ten_j(i)$, and assume that occupational tenure increases productivity at a rate of γ for each additional year of tenure. This leads to the following modified version of Equation (1), where we are explicitly interpreting the potential payoffs in each occupation as wages and therefore denote them as $w_j(i|k)$:

$$w_j(i|k) = p_j f\left[X(i)\right](1 + ten_j(i))^{\gamma}\left(\frac{z_j(i)}{d_{kj}}\right) \tag{A.1}$$

The extra productivity from tenure is due to the accumulation of occupation-specific human capital. It is entirely non-transferable and lost when switching out of occupation j.[2]

With this modified wage specification, the probability that occupation j offers individual i the highest wage, which is the probability that individual i will optimally choose to switch to occupation j, given his current occupation k (denoted by $\pi_{kj}(i)$) is given by:

$$\begin{aligned}
\pi_{kj}(i) &\equiv Pr\left[w_j(i|k) \geq \max_s \{w_s(i|k)\}\right] \\
&= \int_0^{\infty} Pr\left[w_s(i|k) \leq w, \forall s \neq j\right] \cdot dPr\left[w_j(i|k) \leq w\right] \\
&= \frac{T_j d_{kj}^{-\theta}\left[p_j\left(1 + ten_j(i)\right)^{\gamma}\right]^{\theta}}{\sum_{s=1}^{N} T_s d_{ks}^{-\theta}\left[p_s\left(1 + ten_s(i)\right)^{\gamma}\right]^{\theta}}
\end{aligned} \tag{A.2}$$

Note that $ten_j(i) = 0\ \forall j \neq k$. Therefore, $\forall j \neq k$:

$$\pi_{kj}(i) = \frac{T_j d_{kj}^{-\theta} p_j^{\theta}}{\sum_{s\neq k} T_s d_{ks}^{-\theta} p_s^{\theta} + T_k p_k^{\theta}\left(1 + ten_k(i)\right)^{\gamma\theta}} \tag{A.3}$$

Meanwhile, individual i's probability of staying in occupation k, π_{kk} is given by:

[1] For evidence on the importance of occupation-specific human capital, see Kambourov and Manovskii (2009b).

[2] Occupation-specific human capital is assumed to be transferable across employers within the same occupation but is completely lost when switching occupations.

$$\pi_{kk}(i) = \frac{T_k p_k^\theta \left(1 + ten_k(i)\right)^{\gamma\theta}}{\sum_{s \neq k} T_s d_{ks}^{-\theta} p_s^\theta + T_k p_k^\theta \left(1 + ten_k(i)\right)^{\gamma\theta}} \tag{A.4}$$

Dividing (A.3) by (A.4), and taking logs of the ratio, we have:

$$\ln \frac{\pi_{kj}(i)}{\pi_{kk}(i)} = \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k - \theta \ln d_{kj} - \gamma\theta \ln(1 + ten_k(i)) \tag{A.5}$$

Averaging this across individuals in occupation k leads to the gravity-type equation:

$$\begin{aligned} \frac{1}{N_k} \sum_{i=1}^{N_k} \ln \frac{\pi_{kj}(i)}{\pi_{kk}(i)} &= \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k \\ &\quad - \theta \ln d_{kj} - \gamma\theta \frac{1}{N_k} \sum_{i=1}^{N_k} \ln(1 + ten_k(i)) \end{aligned} \tag{A.6}$$

where N_k is the number of individuals in occupation k.

Note that the right-hand-side of the equation is the same as in the main body of the paper, with the addition of a weighted average of log-tenure in the source occupation. Given that the only individual-specific component on the right-hand-side of the equation is occupational tenure, all individuals with tenure level x have the same transition probabilities. With access to a dataset with a large number of individuals at a number of different tenure levels, the left-hand-side of the equation could be empirically measured as a weighted average:

$$\sum_x \frac{N_k^x}{N_k} \ln \frac{sw_{kj}^x}{sw_{kk}^x} \tag{A.7}$$

where N_k^x is the number of individuals in occupation k with tenure level x (at the start of the period), sw_{kj}^x represents the number of switchers from occupation k to occupation j with tenure x, and the sum is over the different levels of x.

However, it can also be shown that:

$$\frac{1}{N_k} \sum_{i=1}^{N_k} \ln \frac{\pi_{kj}(i)}{\pi_{kk}(i)} = \ln \left(\frac{\sum_{i=1}^{N_k} \pi_{kj}(i)}{\sum_{i=1}^{N_k} \pi_{kk}(i)} \right) + c_k \tag{A.8}$$

where c_k is a constant specific to occupation k. Moreover, with a large number of individuals in each occupation we have that:

$$\ln \left(\frac{\sum_{i=1}^{N_k} \pi_{kj}(i)}{\sum_{i=1}^{N_k} \pi_{kk}(i)} \right) = \ln \left(\frac{sw_{kj}}{sw_{kk}} \right) \tag{A.9}$$

Given (A.8) and (A.9) we can rewrite the gravity equation (A.6) as:

$$\ln \frac{sw_{kj}}{sw_{kk}} = \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k + c_k \\ - \theta \ln d_{kj} - \gamma\theta \frac{1}{N_k} \sum_{i=1}^{N_k} \ln(1 + ten_k(i)) \tag{A.10}$$

This can be estimated exactly as in the main text using source and destination occupation fixed effects and a set of proxies for mobility costs. However, the interpretation of the estimated source occupation fixed effects would change, as they would reflect not only T_k and p_k, but also the adjustment factor c_k as well as the effects of occupational tenure. The change in the interpretation of the source fixed effect is similar to what is obtained from the alternative model specification that features exit costs discussed in Appendix B.

Appendix B Alternative Specification: Occupation Exit Costs

Consider an alternative setup featuring occupation-specific exit costs, rather than occupation access costs m_j as in the baseline model. This implies:

$$\ln d_{kj} = \beta_1 dist_{kj} + \beta_2 \lambda_{kj}^{NC} + \beta_3 \lambda_{kj}^{RC} + \beta_4 \lambda_{kj}^{RM} + \beta_5 \lambda_{kj}^{NM} + \chi_k + \epsilon_{kj} \tag{A.11}$$

where χ_k is the cost of leaving occupation k towards any new occupation. Equation (A.11) leads to the following estimating equation:

$$\ln\left(\frac{sw_{kj}}{sw_{kk}}\right) = D_j - S_k - \theta\beta_1 dist_{kj} - \theta\beta_2 \lambda_{kj}^{NC} - \theta\beta_3 \lambda_{kj}^{RC} - \theta\beta_4 \lambda_{kj}^{RM} - \theta\beta_5 \lambda_{kj}^{NM} - \theta\epsilon_{kj} \tag{A.12}$$

where now $S_k \equiv \ln T_k + \theta \ln p_k + \theta\chi_k$, and $D_j \equiv \ln T_j + \theta \ln p_j$. Under this alternative specification, the attractiveness of an occupation is reflected in the *destination* fixed effect. The transition cost is still identified by the difference between the source and destination fixed effect for occupation j (as in the baseline specification in the main body of the paper), given that $\theta\chi_j = S_j - D_j$, but now this cost is interpreted as an *exit* cost rather than an access cost.[3]

[3] Appendix A illustrates how non-transferable occupation-specific skills may act as an exit cost which would be included in the estimated source occupation fixed effect.

More generally, allowing for both entry and exit costs would lead to a specification where:

$$\ln d_{kj} = \beta_1 dist_{kj} + \beta_2 \lambda_{kj}^{NC} + \beta_3 \lambda_{kj}^{RC} + \beta_4 \lambda_{kj}^{RM} + \beta_5 \lambda_{kj}^{NM} + m_j + \chi_k + \epsilon_{kj} \tag{A.13}$$

And therefore:

$$\ln\left(\frac{sw_{kj}}{sw_{kk}}\right) = D_j - S_k - \theta\beta_1 dist_{kj} - \theta\beta_2\lambda_{kj}^{NC} - \theta\beta_3\lambda_{kj}^{RC} - \theta\beta_4\lambda_{kj}^{RM} - \theta\beta_5\lambda_{kj}^{NM} - \theta\epsilon_{kj} \tag{A.14}$$

where $S_k \equiv \ln T_k + \theta \ln p_k + \theta\chi_k$, and $D_j \equiv \ln T_j + \theta \ln p_j - \theta m_j$. The difference between the source and destination fixed effects for occupation j now identifies the sum of the entry and exit costs: $\theta(m_j + \chi_j) = S_j - D_j$. However, we would not be able to separately identify each of these two components.

Both conceptually and in terms of measurement, entry and exit costs are difficult to distinguish. If an occupation requires a large investment of specific human capital that is not valued in other occupations, one can view this as a large entry cost for outsiders. However, from the perspective of insiders this large investment represents an exit cost: they would lose the return to their investment if they switched to another occupation. Hence the fact that an occupation uses a specific set of skills which are not valued in other occupations may create both an entry barrier (to outsiders) and a lock-in effect (to insiders). In general, many of the factors that limit an occupation's accessibility may also make exiting that occupation more difficult.

One way to gain some information about the relative magnitude of entry and exit costs is to examine the variance of the source and destination fixed effects. Let $\widetilde{T}_j \equiv \ln T_j + \theta \ln p_j$, so that $S_j = \widetilde{T}_j + \theta\chi_j$, and $D_j = \widetilde{T}_j - \theta m_j$.

The variance of the source and destination fixed effects are then given by:

$$Var(S_j) = Var\left(\widetilde{T}_j\right) + \theta^2 Var\left(\chi_j\right) + 2\theta Cov\left(\widetilde{T}_j, \chi_j\right)$$
$$Var(D_j) = Var\left(\widetilde{T}_j\right) + \theta^2 Var\left(m_j\right) - 2\theta Cov\left(\widetilde{T}_j, m_j\right)$$

It follows that:

$$Var(S_j) - Var(D_j) = \theta^2\left[Var\left(\chi_j\right) - Var\left(m_j\right)\right] + 2\theta\left[Cov\left(\widetilde{T}_j, \chi_j\right) + Cov\left(\widetilde{T}_j, m_j\right)\right]$$

Under the assumption that entry and exit costs are independent of occupational characteristics captured in $\widetilde{T}_j$, the above expression simplifies to:

$$Var(S_j) - Var(D_j) = \theta^2\left[Var\left(\chi_j\right) - Var\left(m_j\right)\right] \tag{A.15}$$

and thus the difference between the variance of the source and destination fixed effects offers a way to gauge whether the variance of the exit costs is larger or smaller than the variance of the access costs.

Pooling all years and using occupation sizes as weights, the variance of the source and destination fixed effects are found to be as follows:

$$Var(S_j) = 0.295 \qquad\qquad Var(D_j) = 1.124$$

The variance of the destination fixed effects is much larger than that of the source fixed effects, which would imply that the variance of the access costs is larger than the variance of the occupation exit costs.

It is also useful to observe that:

$$\begin{aligned} Var\left(S_j - D_j\right) &= \theta^2 Var\left(m_j + \chi_j\right) \\ &= \theta^2 \left[Var\left(m_j\right) + Var\left(\chi_j\right) + 2Cov\left(m_j, \chi_j\right)\right]. \end{aligned}$$

If we assume that access and exit costs capture different features of a job and are independent of each other, we have that:

$$Var\left(S_j - D_j\right) = \theta^2 \left[Var\left(m_j\right) + Var\left(\chi_j\right)\right]. \tag{A.16}$$

Combining Equations (A.15) and (A.16) yields:

$$Var\left(S_j - D_j\right) - \left[Var(S_j) - Var(D_j)\right] = 2\theta^2 Var(m_j)$$

which, given an estimate of θ, allows us to identify the variance of the occupation access costs, $Var(m_j)$. The variance of the exit costs, $Var(\chi_j)$, can be residually identified using either Equation (A.15) or (A.16).

Using the baseline value $\theta = 3.23$, this yields:

$$Var(m_j) = 0.144 \qquad\qquad Var(\chi_j) = 0.064.$$

These estimates suggest that the variance in access costs is more than twice as large as the variance of occupational exit costs, and therefore occupation access costs account for the majority of the variation in occupation-specific transition costs. Hence, in the main body of the paper we maintain the baseline interpretation that estimated occupation-specific costs primarily reflect access, rather than exit, costs.

Appendix C Matching DOT with CPS

The National Crosswalk Service Center provides a crosswalk between the occupation codes in the 1991 Dictionary of Occupational Titles (DOT) and the 1990 Census Occupation Codes (COC).[4] 1990-COC codes are first converted to the standardized 3-digit occupation codes from Autor and Dorn (2013), which are adapted from Meyer and Osborne (2005). Next, because the DOT classification is much more detailed than the standardized occupation codes, unweighted means are calculated for each DOT dimension at the standardized occupation code level. Each dimension of the DOT is then rescaled to have mean zero and standard deviation one across the universe of standardized occupation codes. Finally, to generate scores at the 2-digit level, an unweighted average is taken across all 3-digit occupations that are within the same 2-digit category.

Appendix D Alternative Ways to Approximate θ

Our main estimation equation is:

$$\ln\left(\frac{sw_{kj}}{sw_{kk}}\right) = D_j - S_k - \theta\beta_1 dist_{kj} - \theta\beta_2\lambda_{kj}^{NC} - \theta\beta_3\lambda_{kj}^{RC} - \theta\beta_4\lambda_{kj}^{RM} - \theta\beta_5\lambda_{kj}^{NM} - \theta\epsilon_{kj}$$

where $S_k \equiv \ln T_k + \theta \ln p_k$ and $D_j \equiv S_j - \theta m_j$.

Empirically, one source and one destination occupation must be excluded from the regression and a normalization must be made. In the main body of the text, we make the normalization $S_2 = 0$ (for occupation code 2, "Executives, administrators and managers"), which implies assuming $T_2 = 1$ and $p_2 = 1$ in each period. We interpret the constant obtained from the regression as the destination effect for the omitted occupation (D_2). Given the definition of D_j and the normalization $S_2 = 0$, this implies that the constant is equal to $-\theta m_2$. By obtaining an estimate of θ following the procedure described in Section 4.3, we can back out a value of m_j for all occupations, including occupation 2.

Alternatively, a further normalization could be made such that $m_2 = 1$ (in addition to $S_2 = 0$). In this case, we can directly interpret the constant from the regression as an estimate of $-\theta$. This additional normalization allows one to obtain an estimate of θ without relying on wage data. Given the estimated constant in Column (3) of Table 2, this normalization would yield an estimate of θ of 3.63 for the year 2012. This estimate is close to the benchmark estimate of 3.23 used in the main body of the text.

Clearly in this case the estimated value of θ will depend on the occupation that is chosen as the omitted category. The estimate that would be obtained for the constant in the year

[4] The crosswalk is the National Occupational Information Coordination Committee (NOICC) Master Crosswalk, Version 4.3, downloadable from `ftp://ftp.xwalkcenter.org/download/xwalks/`, file `xwalkv43.exe`.

2012 when a particular occupation is omitted can be inferred directly from the y-axis in Figure 1. In this case omitting occupation 2 happens to yield the lowest estimate of θ (3.63), while omitting occupation 12 would yield the highest estimate (8.87). In Table 9 we consider the robustness of our results to the full range of possible values of θ that would be obtained from this approach.

Appendix E Additional Evidence on Non-Pecuniary Returns

As discussed in Section 4.3, the baseline estimate of match quality dispersion could be positively or negatively biased depending on the sign and intensity of the covariation between current wages and other components of total lifetime utility payoffs, including non-pecuniary returns. This covariation cannot be directly approximated using CPS data. For this reason in Section 5.2 we perform multiple robustness checks of our results, setting widely different values for match quality dispersion.

In what follows we present evidence on the covariation between pecuniary and non-pecuniary returns, and examine the relative dispersion of self-reported non-pecuniary rewards. To this purpose we resort to information from alternative data sets containing proxies of non-pecuniary returns (job satisfaction measures) and current wages. We use data from two surveys administered by the US National Science Foundation: the 2010 National Survey of Recent College Graduates (NSRCG) and the 2013 National Survey of College Graduates (NSCG).[5]

The National Survey of College Graduates (NSCG) is sponsored by the National Center for Science and Engineering Statistics (NCSES) at the NSF. The Census Bureau is responsible for data collection. The survey provides data on a number of characteristics of individuals with a bachelor's or higher degree, with a special focus on individuals with education and/or employment in science or engineering. The National Survey of Recent College Graduates is similar and also provides information about individuals holding a bachelor's or master's degree in a science, engineering, or health field from a U.S. academic institution.

Both surveys are cross-sectional and, crucially, they contain information about salary and job satisfaction of sample members. The (self-reported) job satisfaction measures reflect different aspects of match quality in the current occupation. Table A.5 summarizes the different satisfaction measures and shows the specific job features they capture.

Satisfaction is measured on a four-point scale. We convert the responses so that they are increasing in satisfaction and work with logarithms to focus on proportional variation. The survey provides information about the annual salary of the respondent.[6] Focusing on workers

[5]Information about these data can be found at http://www.nsf.gov/statistics/sestat/.

[6]The survey question asks: "what was your basic annual salary on your principal job, before deductions?"

who report that they were employed over the whole year, we generate an implied hourly wage rate using information on weekly hours of work. We restrict the sample to workers up to age 65 who report working between 20 and 84 hours per week. We exclude workers with salaries below \$6,000 or above \$400,000 per year. We then construct residual wages by running a regression of log hourly wages on a quartic in age, a female dummy and interactions of these variables. The resulting samples (featuring non-missing occupation and job satisfaction values) consist of 49,675 observations in the 2010 NSRCG and 69,451 observations in the 2013 NSCG.

We use these data to: (i) gauge the relative dispersion of non-pecuniary returns and contrast it to the dispersion of pecuniary returns, and (ii) compute direct measures of the covariance between pecuniary returns and different satisfaction scores, some of which clearly focus on non-pecuniary aspects of the job. We compute these measures across the full sample, and also conditional on individuals' current occupation.[7]

Table A.6 reports the standard deviation of residual wages and of alternative measures of job satisfaction in each year. Both the unconditional measure and the median measure conditional on occupation are reported.[8] Table A.7 shows the covariance between salary and job satisfaction measures.

We observe that the standard deviation in pecuniary returns is roughly three times larger than its counterpart for non-pecuniary returns (depending on which measure of job satisfaction is considered). Moreover, with only one exception, the covariance between pecuniary and non-pecuniary rewards is positive in both data samples and across different satisfaction measures. The covariances are fairly low and often close to zero.

This additional evidence suggests that estimates of match quality dispersion based on wage dispersion may be lower than the true underlying value, but not grossly so. Focusing on the conditional measures for 2010, the highest measured covariances with residual wages (0.024 and 0.017) are detected for measures of satisfaction about salary and benefits; hence they do not seem appropriate to gauge non-pecuniary aspects of the returns. The next highest covariance is 0.011 and refers to satisfaction about job security. We use this value, along with the corresponding standard deviation from Table A.6 (0.17, which is a typical value for the standard deviation among satisfaction measures) to compute an approximate measure of the extent to which the dispersion measure based on wages alone will underestimate the true value of θ.

[7]Recall that in the main body of the paper we focus on the ex-post dispersions of wages in the CPS, conditional on initial occupations. Unfortunately the NSRCG and NSCG datasets do not allow us to track workers over time at a similar frequency as the CPS, so we compute dispersion measures based on current occupation. Occupations are aggregated to the 2-digit level. The coding system used in the NSRCG and NSCG is not the same as in the CPS, but by using 2-digit occupations we achieve a similar level of aggregation to what we use in the main body of the paper.

[8]We exclude occupations with fewer than 100 observations.

Recall that, as we discuss in Section 4.3:

$$Var(\ln \phi) = Var(\ln w) + Var(\ln \varphi) + 2Cov(\ln w, \ln \varphi)$$

The dispersion measure of interest σ, which is the dispersion of total log payoffs, is therefore:

$$\sigma = \sqrt{Var(\ln w) + Var(\ln \varphi) + 2Cov(\ln w, \ln \varphi)}$$

Taking the dispersion of wages as 0.44, the dispersion of satisfaction as 0.17 and the covariance as 0.011, we have that in this dataset $\sigma = 0.4945$.

Recall also that:

$$\theta = \frac{\pi}{\sigma\sqrt{6}}$$

This would yield an estimate of θ based on the NSRCG data of 2.59. Let $\widetilde{\theta}$ denote the estimate of θ based on the standard deviation measured from pecuniary payoffs only. In the NSRCG data we would have $\widetilde{\theta} = 2.91$. This implies $\widetilde{\theta} = 1.12 \cdot \theta$.

The results above suggest that an estimate of θ based on wages alone may overestimate the true value of θ due to the omission of non-pecuniary returns by approximately 12%. Using this as an adjustment factor on our baseline estimate of $\widetilde{\theta} = 3.23$ would yield an implied value of $\theta = 2.87$. This value is well within the range of alternative values we consider for θ in Table 9 in the robustness section.

Appendix F Alternative Measures of Task Content

This Appendix provides details on the robustness exercises discussed in Section 5.1 of the paper, which use alternative measures for the construction of task distance. We first consider an alternative task distance measure which includes additional dimensions from the DOT. The additional DOT dimensions are listed in Appendix Table A.2. Results from the estimation of the gravity equation when these additional dimensions are included are presented in Column (1) of Appendix Table A.8. The outcomes of counterfactual experiments analogous to those in Table 8 are presented in Columns (1) and (2) of Table A.9. The fraction of the transition costs that can be attributed to the task variables is slightly higher using this distance measure, but remains below 14% for the median occupation.

We next construct distance measures based on O*Net, the successor to the DOT. We consider two subsets of data from O*Net Version 14.0 (2009): Work Activities and Skills. The full set of work activities and skills from O*Net are listed in Appendix Tables A.3 and A.4. The results from the counterfactual experiments using work activities as the dimensions included in the construction of task distance are presented in Columns (3) and (4) of Table A.9, while Columns (5) and (6) present the results when the skill dimensions are used. The outcomes are similar, with task distance accounting for around 8.5% of transition costs for the median occupation, and costs associated with transitions across broad task groups accounting for an additional 5 to 7 percentage points.

Columns (7) and (8) of Table A.9 show the results when included a cubic function of distance in our gravity equation estimation. Task-related barriers account for around 13% of transition costs for the median occupation. Column (5) of Appendix Table A.8 shows the results from an alternative specification where we allow the transition costs between different broad task groups to vary with both source and destination. To avoid multicollinearity, we must omit transitions from one broad task group to itself, and transitions from non-routine cognitive occupations to any other task group. Results from the counterfactual experiments using this specification (Columns (9) and (10) of Table A.9) show that task-related barriers still account for around 10% of overall transition costs for the median occupation.

Appendix G Transitions through Unemployment

Our analysis focuses on occupational transitions that occur over consecutive months of employment. Naturally, some occupational transitions may instead involve an intervening period of unemployment (or inactivity). To account for these types of transitions, we analyze occupational flows occurring over a longer time horizon. Specifically, we use our matched dataset to compute occupational flows occurring over 12-month horizons. For any given month, we compute flows of workers between occupation pairs over the period between month m in year t and month m in year $t+1$. This effectively allows us to consider all occupational switches occurring over this period, including those that involve an intervening period of non-employment.[9]

Table A.10 shows the results for the relative importance of tasks as a fraction of transition costs based on transitions over 12-month horizons. The main finding from this exercise is that, when allowing for a longer adjustment period (including a potential intervening period of unemployment), overall estimated costs appear to be lower. We also find that the relative importance of task-related costs is higher, accounting for nearly one fifth of total costs for the median occupation.

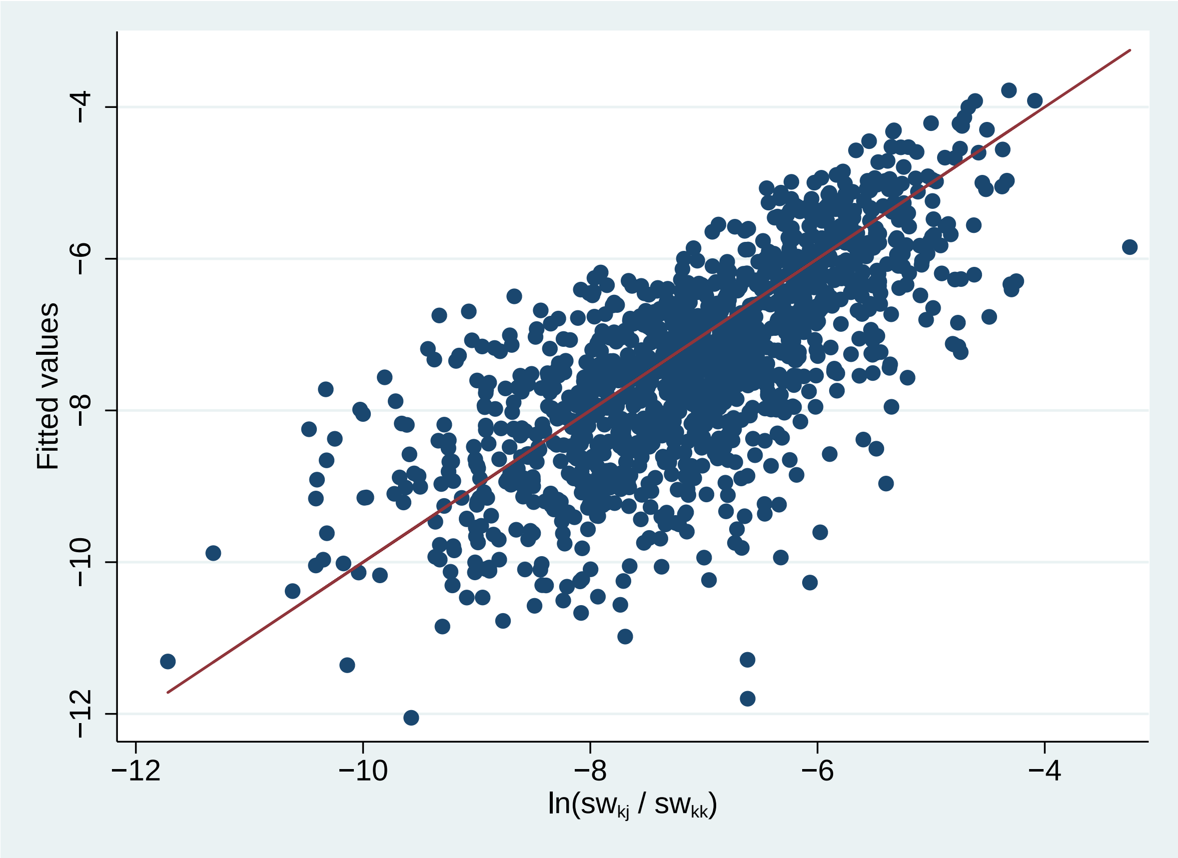
One possible interpretation of these results is that, over longer periods, tasks may play a more important role in occupation mobility decisions. However, a necessary caveat when analyzing these differences is the possibility of bias due to occupation mis-coding. As shown in Section 5.4, coding error leads to an over-estimation of the importance of task content. The

[9] To avoid counting the same individual transition multiple times, we consider only people who are in the outgoing rotation groups (month-in-sample 4 or 8). An alternative approach to account for intervening periods of unemployment would be to consider flows from unemployment to employment based on the previous occupation of the unemployed and their first occupation after unemployment. However, unemployment-to-employment flows at this occupation-pair level are extremely small, making identification infeasible.

main advantage of focusing on month-to-month transitions is the lower prevalence of coding error in the post-1994 period, due to the use of dependent coding techniques. Unfortunately these techniques do not apply when considering transitions at 12-month horizons.[10]

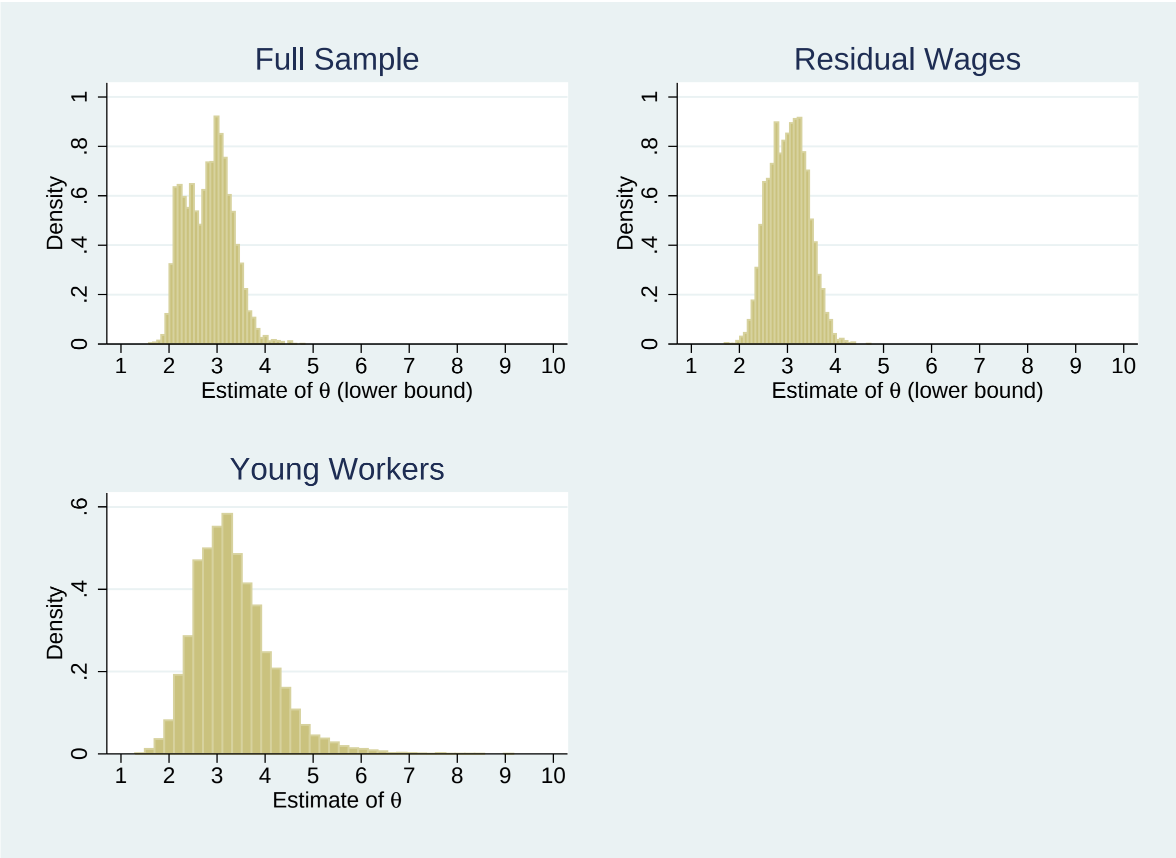
[10]This is due to the rotating nature of the CPS sample. Households are surveyed for four consecutive months, then leave the sample for eights months, and subsequently return for another four months. When households return to the sample for the second four-month spell, they are always independently coded. Moreover, dependent coding techniques do not apply when workers transition to employment from unemployment.

Figure A.1: Goodness of fit



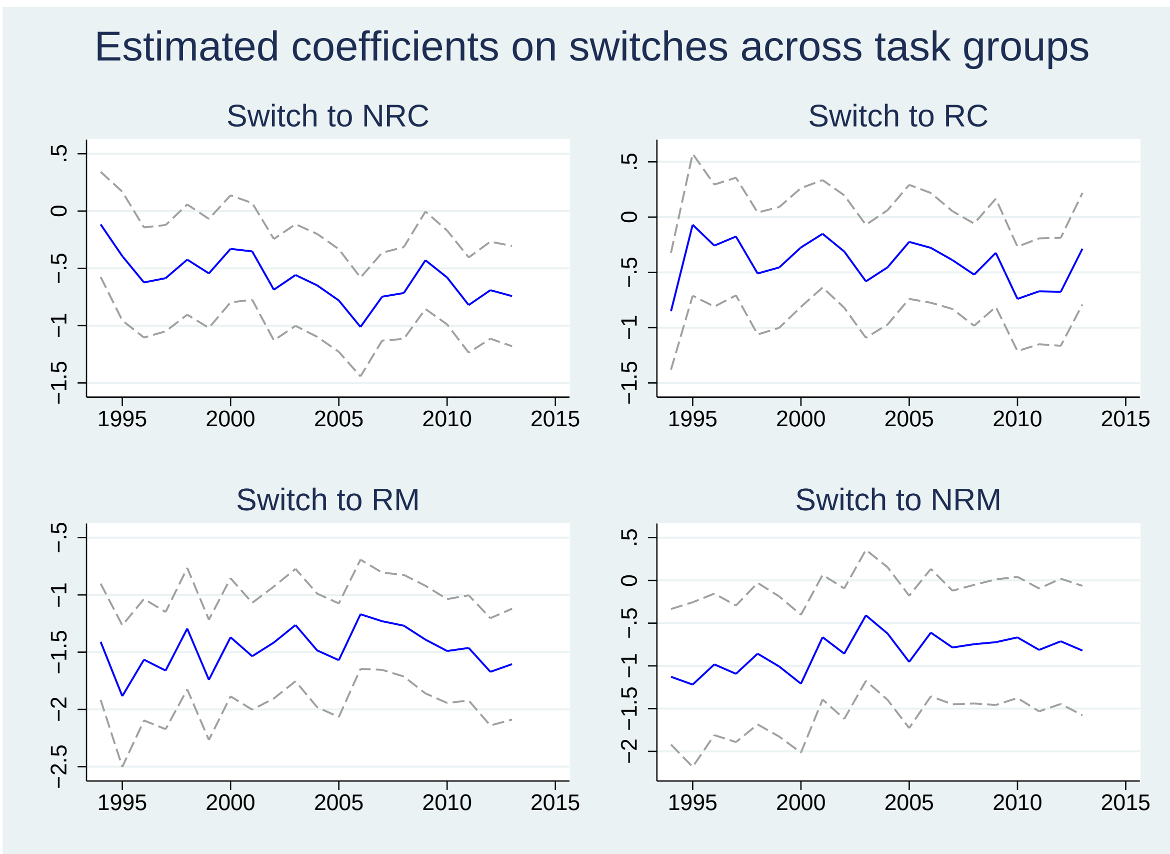
Note: The figure plots the fitted values of the dependent variable $\ln(sw_{kj}/sw_{kk})$ against their true values, based on the estimation in Column (3) of Table 2.

Figure A.2: Histogram of estimated values of θ



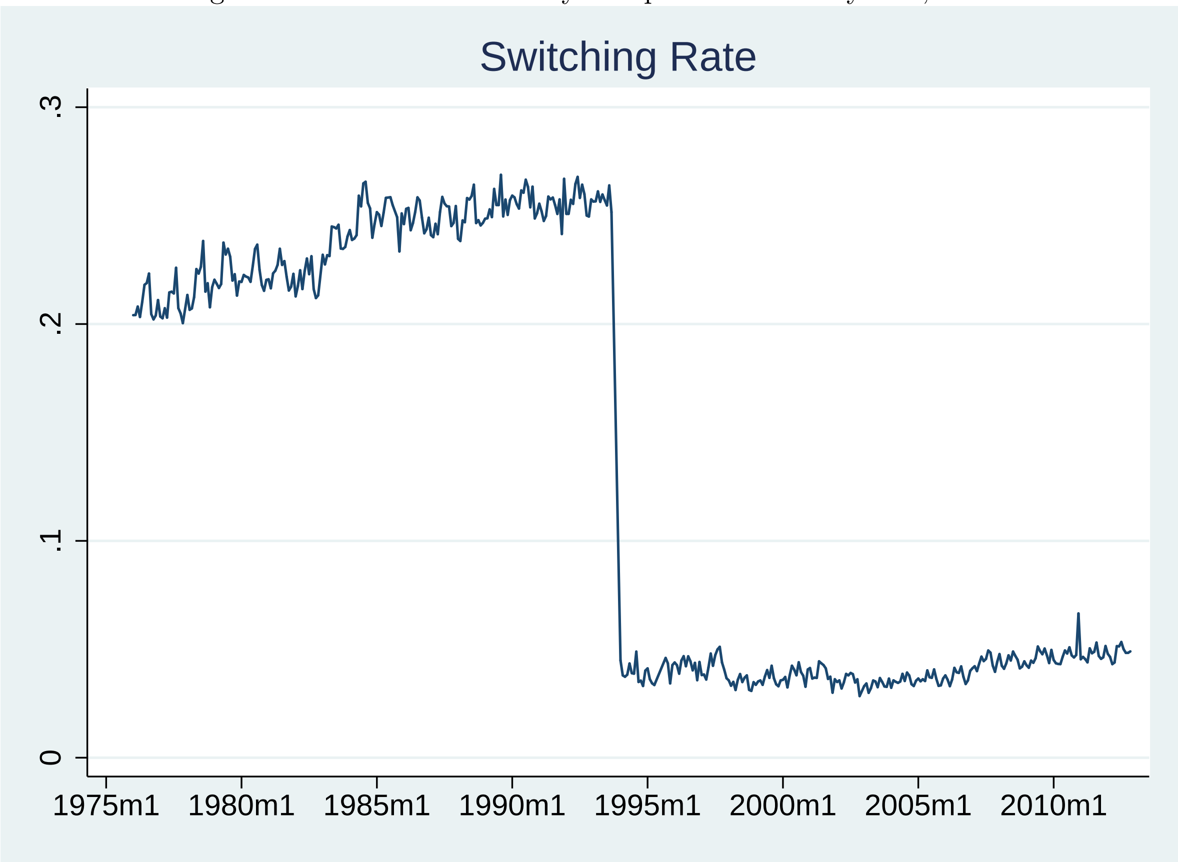
Note: The top panels use occupation-month cells with at least 100 observations; the bottom panel is based on young workers only and uses occupation-gender-month cells with at least 15 observations.

Figure A.3: Evolution of the estimated coefficient on task switching variables over time



Note: The dashed lines indicate 95% confidence intervals.

Figure A.4: Measured monthly occupational mobility rate, CPS



Note: The figure illustrates the discontinuity in measured occupational mobility rates that occurs when dependent coding techniques are introduced in 1994.

Table A.1: 2-digit occupation groupings for the Autor and Dorn (2013) coding system, organized by task categories

2-digit Category	2-digit Code	3-digit Autor and Dorn (2013) Codes
Non-Routine Cognitive:		
Executives, administrators and managers	02	004-022
Management related occupations	03	023-037
Engineers and architects	04	043-059
Mathematical, computer and natural scientists	05	064-083
Health diagnosing occupations	07	084-089
Health assessment and treating occupations	08	095-106
Teachers, college and university	09	154
Teachers, except college and university	10	155-163
Librarians, social scientists, religious workers	11	164-177
Lawyers and judges	12	178
Writers, artists, entertainers, athletes	13	183-199
Health technologists and technicians	14	203-208
Engineering and science technicians	15	214-225
Technicians, except health engineering, and science	16	226-235
Protective service occupations	27	415-427
Routine Cognitive:		
Sales supervisors and sales reps, finance and business	17	243-256
Retail and other salespersons	18	258-283
Office supervisors and computer operators	19	303-308
Secretaries, stenographers, and typists	20	313-315
Information and records processing, except financial	21	316-336
Financial records processing occupations	22	337-344
Office machine operators and mail distributing	24	346-357
Other administrative support occupations, including clerical	25	359-389
Non-Routine Manual:		
Private household cleaners and servers	26	405-408
Food service occupations	28	433-444
Health service occupations	29	445-447
Cleaning and building service occupations, except household	30	448-455
Other personal service occupations	31	457-472
Routine Manual:		
Mechanics and repairers	35	503-549
Construction trades	36	558-599
Other precision production occupations	37	614-699
Machine operators and tenders, not precision	38	703-779
Fabricators, assemblers and hand working occupations	39	783-789
Production inspectors and graders	40	799
Transportation and material moving	41	803-859
Helpers, construction and production occupations	43	865-873
Freight, stock and material handlers	44	875-889

Table A.2: Additional Dimensions, DOT 1991

Temperaments:	
Direction, control, or planning	Performing under stress
Repetitive work	Deal with set limits, tolerances, standards
Influence people	Work under specific instructions
Expressing feelings, ideas, facts	Dealing with people beyond instructions
Variety of duties, often changing	Judgments and decisions
Working alone or in isolation	

Table A.3: List of ONet 2009 Work Activities

4.A.1.a.1	Getting Information
4.A.1.a.2	Monitor Processes, Materials, or Surroundings
4.A.1.b.1	Identifying Objects, Actions, and Events
4.A.1.b.2	Inspecting Equipment, Structures, or Material
4.A.1.b.3	Estimating the Quantifiable Characteristics of Products, Events, or Information
4.A.2.a.1	Judging the Qualities of Things, Services, or People
4.A.2.a.2	Processing Information
4.A.2.a.3	Evaluating Information to Determine Compliance with Standards
4.A.2.a.4	Analyzing Data or Information
4.A.2.b.1	Making Decisions and Solving Problems
4.A.2.b.2	Thinking Creatively
4.A.2.b.3	Updating and Using Relevant Knowledge
4.A.2.b.4	Developing Objectives and Strategies
4.A.2.b.5	Scheduling Work and Activities
4.A.2.b.6	Organizing, Planning, and Prioritizing Work
4.A.3.a.1	Performing General Physical Activities
4.A.3.a.2	Handling and Moving Objects
4.A.3.a.3	Controlling Machines and Processes
4.A.3.a.4	Operating Vehicles, Mechanized Devices, or Equipment
4.A.3.b.1	Interacting With Computers
4.A.3.b.2	Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment
4.A.3.b.4	Repairing and Maintaining Mechanical Equipment
4.A.3.b.5	Repairing and Maintaining Electronic Equipment
4.A.3.b.6	Documenting/Recording Information
4.A.4.a.1	Interpreting the Meaning of Information for Others
4.A.4.a.2	Communicating with Supervisors, Peers, or Subordinates
4.A.4.a.3	Communicating with Persons Outside Organization
4.A.4.a.4	Establishing and Maintaining Interpersonal Relationships
4.A.4.a.5	Assisting and Caring for Others
4.A.4.a.6	Selling or Influencing Others
4.A.4.a.7	Resolving Conflicts and Negotiating with Others
4.A.4.a.8	Performing for or Working Directly with the Public
4.A.4.b.1	Coordinating the Work and Activities of Others
4.A.4.b.2	Developing and Building Teams
4.A.4.b.3	Training and Teaching Others
4.A.4.b.4	Guiding, Directing, and Motivating Subordinates
4.A.4.b.5	Coaching and Developing Others
4.A.4.b.6	Provide Consultation and Advice to Others
4.A.4.c.1	Performing Administrative Activities
4.A.4.c.2	Staffing Organizational Units
4.A.4.c.3	Monitoring and Controlling Resources

Table A.4: List of ONet 2009 Skills

2.A.1.a	Reading Compreh	2.B.3.b	Technology Design
2.A.1.b	Active Listening	2.B.3.c	Equipment Selection
2.A.1.c	Writing	2.B.3.d	Installation
2.A.1.d	Speaking	2.B.3.e	Programming
2.A.1.e	Mathematics	2.B.3.g	Operation Monitoring
2.A.1.f	Science	2.B.3.h	Operation and Control
2.A.2.a	Critical Thinking	2.B.3.j	Equipment Maintenance
2.A.2.b	Active Learning	2.B.3.k	Troubleshooting
2.A.2.c	Learning Strategies	2.B.3.l	Repairing
2.A.2.d	Monitoring	2.B.3.m	Quality Control Analysis
2.B.1.a	Social Perceptiveness	2.B.4.e	Judgment and Decision Mkg
2.B.1.b	Coordination	2.B.4.g	Systems Analysis
2.B.1.c	Persuasion	2.B.4.h	Systems Evaluation
2.B.1.d	Negotiation	2.B.5.a	Time Management
2.B.1.e	Instructing	2.B.5.b	Mgmnt of Financial Resources
2.B.1.f	Service Orientation	2.B.5.c	Mgmnt of Material Resources
2.B.2.i	Complex Problem Solv	2.B.5.d	Mgmnt of Personnel Resources
2.B.3.a	Operations Analysis		

Table A.5: List of job satisfaction measures in SESTAT (NCSG and NSRCG surveys)

Variable name	Area	Questionnaire question
		"Thinking about your principal job, please rate:"
JOBSATIS	Overall	your overall satisfaction
SATADV	Advancement	your satisfaction with that job's opportunities for advancement
SATBEN	Benefits	your satisfaction with that job's benefits
SATCHAL	Challenge	your satisfaction with that job's intellectual challenge
SATIND	Independence	your satisfaction with that job's degree of independence
SATLOC	Location	your satisfaction with that job's job location
SATRESP	Responsibility	your satisfaction with that job's level of responsibility
SATSAL	Salary	your satisfaction with that job's salary
SATSEC	Security	your satisfaction with that job's job security
SATSOC	Contribution	your satisfaction with that job's contribution to society

Table A.6: Standard deviation of residual wages and job satisfaction measures in SESTAT (all logarithms)

	Standard Deviation, 2010		Standard Deviation, 2013	
	Unconditional	Conditional (median)	Unconditional	Conditional (median)
Residual Wage	.55	.44	.56	.46
Job satisfaction:				
Overall	.15	.14	.16	.14
Advancement	.21	.19	.21	.20
Benefits	.20	.17	.20	.18
Challenge	.18	.16	.19	.16
Independence	.15	.14	.15	.14
Location	.15	.16	.16	.16
Responsibility	.16	.14	.16	.15
Salary	.18	.17	.19	.18
Security	.18	.17	.18	.18
Contribution	.18	.15	.18	.15

Note: Results based on NSRCG 2010 and NCSG 2013 surveys. Conditional estimates report the value for the median occupation.

Table A.7: Covariance of residual wages and job satisfaction measures in SESTAT (all logarithms)

	Covariance with residual wage, 2010		Covariance with residual wage, 2013	
	Unconditional	Conditional (median)	Unconditional	Conditional (median)
Job satisfaction:				
Overall	.016	.008	.020	.007
Advancement	.019	.009	.022	.004
Benefits	.030	.017	.035	.020
Challenge	.021	.005	.023	.003
Independence	.008	.003	.010	.003
Location	.001	.001	.003	.002
Responsibility	.013	.003	.015	.002
Salary	.039	.024	.045	.027
Security	.016	.011	.016	.010
Contribution	.006	.001	.006	-.001

Note: Results based on NSRCG 2010 and NCSG 2013 surveys. Conditional estimates report the value for the median occupation.

Table A.8: Robustness checks using alternative task measures

	DOT Alternative	O*Net Work Activ	O*Net Skills	Benchmark Non-linear	Benchmark Task Pairs
	(1)	(2)	(3)	(4)	(5)
$dist$	-2.027	-2.497	-2.526	-1.894	-1.385
	(.241)***	(.209)***	(.201)***	(1.520)	(.213)***
$dist^2$				1.023	
				(3.592)	
$dist^3$				-.650	
				(2.433)	
λ^{NC}	-.566	-.466	-.833	-.698	
	(.217)***	(.197)**	(.181)***	(.218)***	
λ^{RC}	-.807	-.557	-.567	-.672	
	(.248)***	(.242)**	(.240)**	(.263)**	
λ^{RM}	-1.481	-.937	-.796	-1.673	
	(.243)***	(.246)***	(.247)***	(.239)***	
λ^{NM}	-.721	-.572	-.949	-.706	
	(.372)*	(.362)	(.359)***	(.375)*	
$\lambda^{RC\Rightarrow NC}$					-1.339
					(.311)***
$\lambda^{RC\Rightarrow RM}$					-.586
					(.320)*
$\lambda^{RC\Rightarrow NM}$					-1.033
					(.374)***
$\lambda^{RM\Rightarrow NC}$					-2.506
					(.354)***
$\lambda^{RM\Rightarrow RC}$					-1.789
					(.332)***
$\lambda^{RM\Rightarrow NM}$					-1.599
					(.351)***
$\lambda^{NM\Rightarrow NC}$					-1.464
					(.478)***
$\lambda^{NM\Rightarrow RC}$					-.605
					(.468)
$\lambda^{NM\Rightarrow RM}$					-.620
					(.448)
Const.	-3.397	-3.255	-3.172	-3.596	-3.645
	(.358)***	(.347)***	(.346)***	(.382)***	(.356)***
Obs.	1332	1332	1332	1332	1332

Note: The table presents the results from the estimation of Equation (15) for the year 2012 using alternative task measures. The dependent variable is $\ln(sw_{kj}/sw_{kk})$. All specifications include source and destination occupation dummies.

Table A.9: Robustness checks for the results from the counterfactual experiments using alternative task measures

	DOT - Alternative		O*Net Work Act		O*Net Skills		Cubic		Task Pairs	
	Distance Fraction	Tasks Fraction	Distance Fraction	Tasks Fraction	Distance Fraction	Tasks Fraction	Distance Fraction	Tasks Fraction	Distance Fraction	Tasks Fraction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
10th Percentile	.036	.074	.043	.068	.039	.071	.031	.066	.023	.032
25th Percentile	.051	.103	.063	.102	.060	.110	.043	.094	.031	.047
50th Percentile	.069	.138	.085	.135	.082	.151	.058	.125	.040	.095
75th Percentile	.088	.175	.107	.171	.104	.195	.076	.161	.052	.187
90th Percentile	.109	.211	.129	.206	.127	.231	.096	.201	.063	.249
Maximum	.182	.344	.210	.335	.223	.360	.178	.324	.135	.459
Mean	.071	.142	.086	.137	.084	.152	.061	.131	.042	.121
Obs.	740	740	740	740	740	740	740	740	740	740

Note: The observations are occupation-year cells. Columns (1), (3), (5), (7) and (9) present the summary statistics for the fraction of the counterfactual mobility increase that can be attributed to task distance using the specification indicated on the first row. Columns (2), (4), (6), (8) and (10) present the fraction that can be attributed to all task-related barriers (task distance and costs of transitioning across broad task groups). The remainder is accounted for by heterogeneity in task-independent occupational entry costs.

Table A.10: Summary statistics for the relative size of the transition cost associated with the task-related variables; estimation based on worker flows over 12-month time horizon

	Distance	Tasks
	(1)	(2)
10th Percentile	0.012	0.024
25th Percentile	0.039	0.077
50th Percentile	0.084	0.180
75th Percentile	0.138	0.303
90th Percentile	0.217	0.462
Mean	0.141	0.255
Obs.	23,976	23,976

Note: The observations are occupation pair-year cells. Column (1) presents the summary statistics for the fraction of the transition costs that can be attributed to task distance, while Column (2) presents the fraction that can be attributed to all task-related barriers (task distance and costs of transitioning across broad task groups). The remainder is accounted for by task-independent occupational entry costs.