ONLINE APPENDIX FOR "THE COSTS OF OCCUPATIONAL MOBILITY: AN AGGREGATE ANALYSIS"

Guido Matias Cortes
University of Manchester

Giovanni Gallipoli
University of British Columbia

The editor in charge of this paper was Dirk Krueger.

E-mail: matias.cortes@manchester.ac.uk (Cortes); gallipol@mail.ubc.ca (Gallipoli)

Journal of the European Economic Association
Preprint prepared on 24 November 2016 using jeea.cls *v1.0.*

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:

$$\ln d_{kj} = \beta_1 dist_{kj} + \beta_2 \lambda_{kj}^{NC} + \beta_3 \lambda_{kj}^{RC} + \beta_4 \lambda_{kj}^{RM} + \beta_5 \lambda_{kj}^{NM} + \chi_k + \epsilon_{kj} \tag{B.1}$$

where χ_k is the cost of leaving occupation k towards any new occupation. Equation (B.1) leads to the following estimating equation:

$$\ln \left(\frac{sw_{kj}}{sw_{kk}} \right) = D_j - S_k - \theta\beta_1 dist_{kj} - \theta\beta_2 \lambda_{kj}^{NC} - \theta\beta_3 \lambda_{kj}^{RC} - \theta\beta_4 \lambda_{kj}^{RM} - \theta\beta_5 \lambda_{kj}^{NM} - \theta\epsilon_{kj} \tag{B.2}$$

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:

$$\ln d_{kj} = \beta_1 dist_{kj} + \beta_2 \lambda_{kj}^{NC} + \beta_3 \lambda_{kj}^{RC} + \beta_4 \lambda_{kj}^{RM} + \beta_5 \lambda_{kj}^{NM} + m_j + \chi_k + \epsilon_{kj} \tag{B.3}$$

And therefore:

$$\ln\left(\frac{sw_{kj}}{sw_{kk}}\right) = D_j - S_k - \theta\beta_1 dist_{kj} - \theta\beta_2 \lambda_{kj}^{NC} - \theta\beta_3 \lambda_{kj}^{RC} - \theta\beta_4 \lambda_{kj}^{RM} - \theta\beta_5 \lambda_{kj}^{NM} - \theta\epsilon_{kj} \tag{B.4}$$

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:

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

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:

$$Var(S_j) - Var(D_j) = \theta^2 \left[Var\left(\chi_j\right) - Var\left(m_j\right)\right] \tag{B.5}$$

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:

$$Var\left(S_j - D_j\right) = \theta^2\left[Var\left(m_j\right) + Var\left(\chi_j\right)\right]. \tag{B.6}$$

Combining Equations (B.5) and (B.6) 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 (B.5) or (B.6).

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 E.1 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 E.2 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 E.3 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 E.2 (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:

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

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

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

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 F.1. The outcomes of counterfactual experiments analogous to those in Table 8 are presented in Columns (1) and (2) of Table F.2. 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 F.2, 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 F.2 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 F.1 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 F.2) 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 G.1 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.

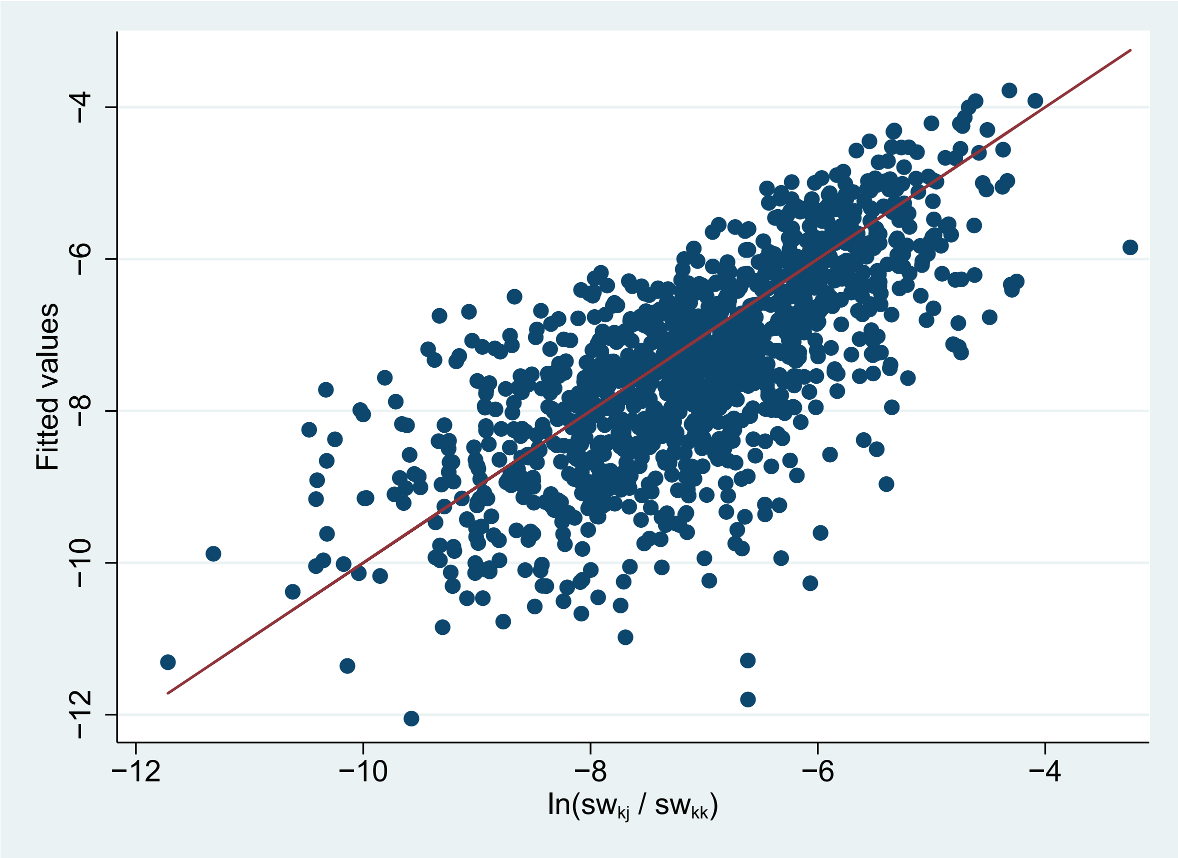
One possible interpretation of these results is that, over longer periods, tasks may play a more important role in occupation mobility decisions. However, a necessary caveat when analyzing these differences is the possibility of bias due to occupation mis-coding. As shown in Section 5.4, coding error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

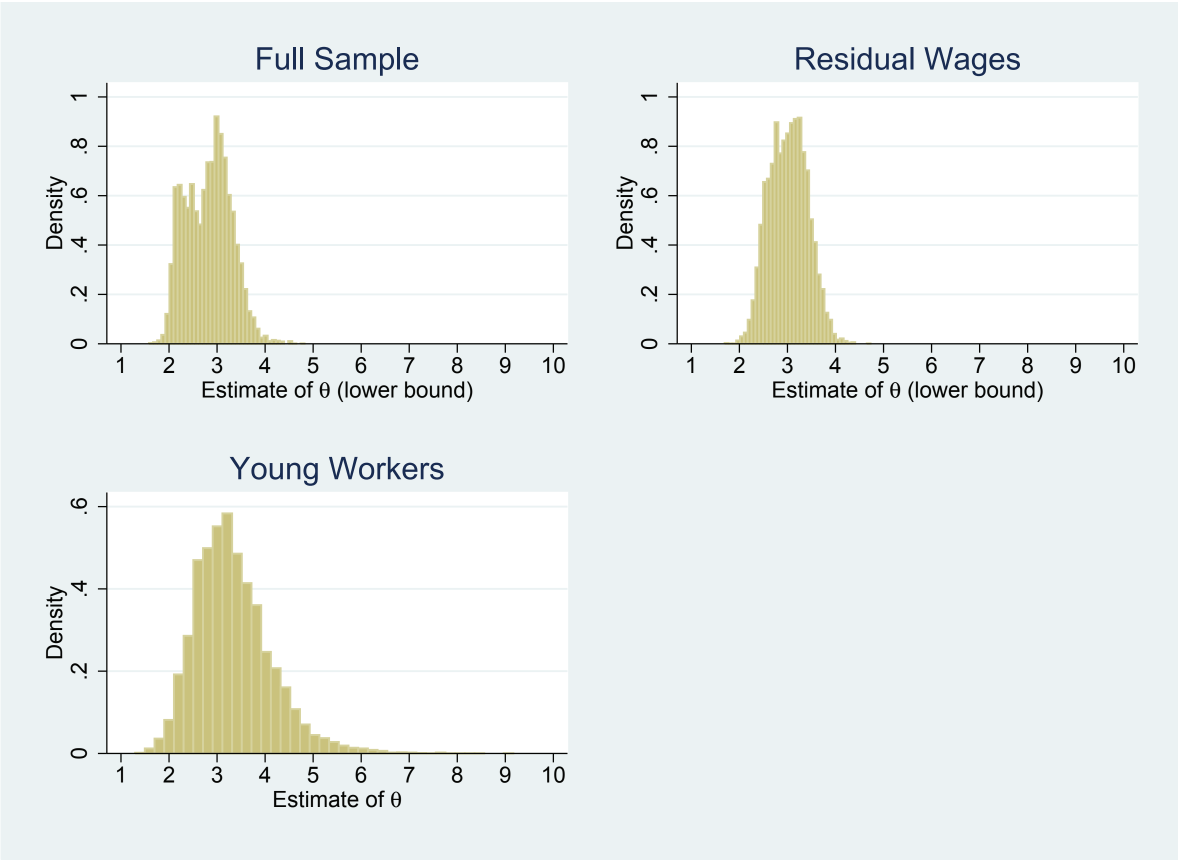
[10]This is due to the rotating nature of the CPS sample. Households are surveyed for four consecutive months, then leave the sample for eights months, and subsequently return for another four months. When households return to the sample for the second four-month spell, they are always independently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

Figure A.1: Goodness of fit



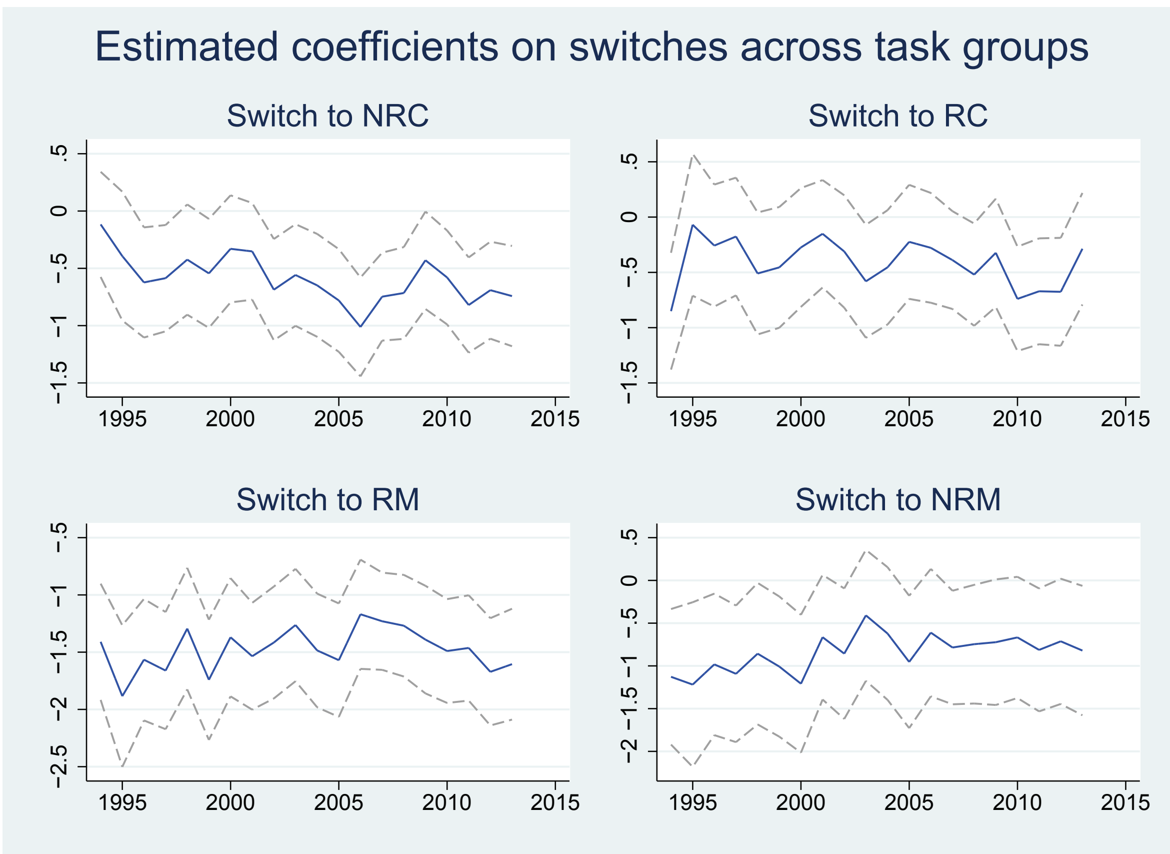
Note: The figure plots the fitted values of the dependent variable $\ln(sw_{kj}/sw_{kk})$ against their true values, based on the estimation in Column (3) of Table 2.

Figure A.2: Histogram of estimated values of θ



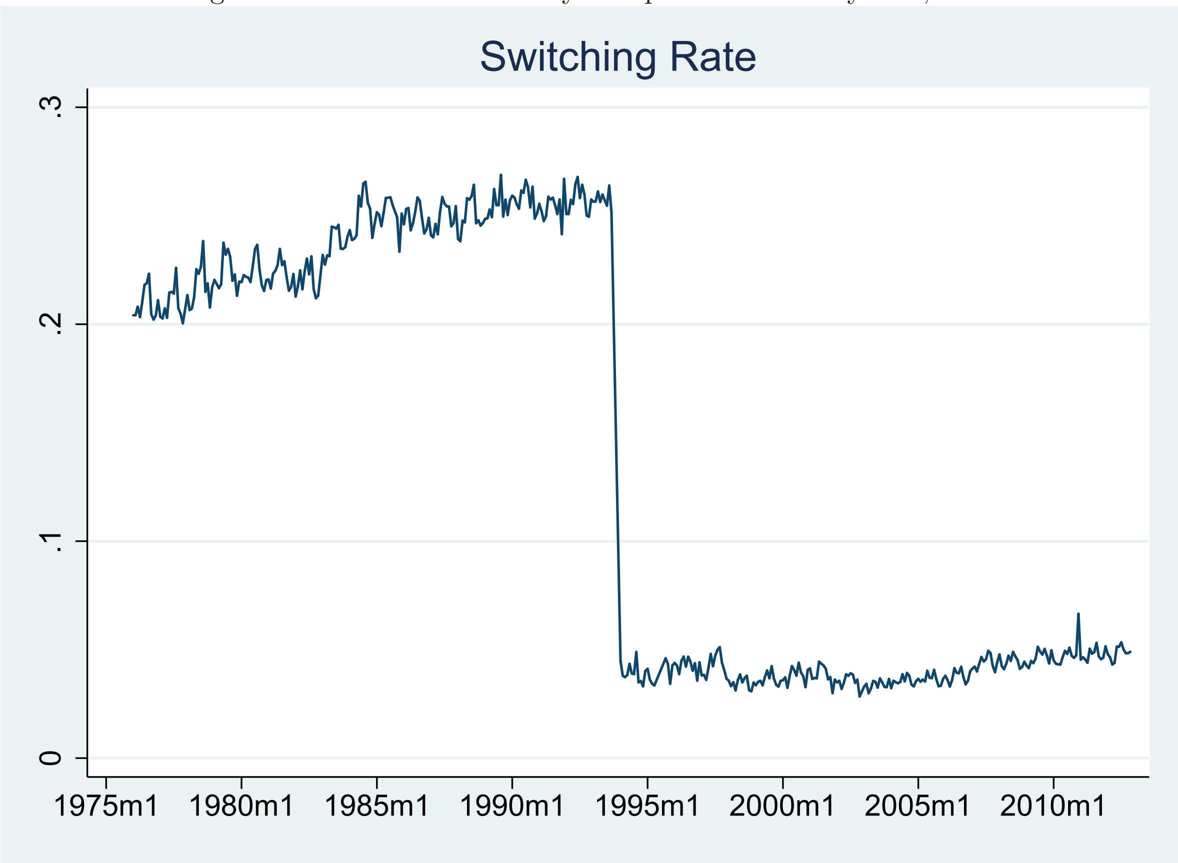
Note: The top panels use occupation-month cells with at least 100 observations; the bottom panel is based on young workers only and uses occupation-gender-month cells with at least 15 observations.

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 E.1: 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 E.2: 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 E.3: 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 F.1: 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 F.2: 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 G.1: 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.