

Federal Reserve Bank of Minneapolis
Research Department

Understanding the Long-Run Decline in Interstate Migration: Online Appendix*

Greg Kaplan and Sam Schulhofer-Wohl

Working Paper 725

June 2015

ABSTRACT

This appendix contains eight sections. Section 1 gives technical details of how we calculate standard errors in the CPS data. Section 2 discusses changes in the ACS procedures before 2005. Section 3 examines demographic and economic patterns in migration over the past two decades, in more detail than in the main paper. Section 4 examines the cross-sectional variance of location-occupation interactions in earnings when we define locations by MSAs instead of states. Section 5 describes alternative methods to estimate the variance of location-occupation interactions in income. Section 6 measures the segregation of industries across states and of occupations and industries across MSAs. Section 7 gives technical details on the use of SIPP and census data to calculate repeat and return migration rates. Section 8 discusses transition dynamics in the model.

*Kaplan: Princeton University and National Bureau of Economic Research; gkaplan@princeton.edu. Schulhofer-Wohl: Federal Reserve Bank of Minneapolis; wohls@minneapolisfed.org. We thank Diego Amador, Chloe Booth, and Yoon Sun Hur for excellent research assistance and Joan Gieseke for editorial assistance. We also thank numerous seminar and conference participants for helpful comments. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Standard error calculation in CPS data

From 2005 onward, we calculate standard errors using the person-level replicate weights provided by the Census Bureau that account for the design of the Current Population Survey (CPS) sample. Before 2005, replicate weights are not available, so we calculate standard errors by assuming that the survey weights are inversely proportional to the probability of sampling and that the sample is clustered by households. Clustering on households and replicate weights give virtually identical standard errors for the interstate migration rate in 2005 and later years. We do not follow Davern et al.'s (2006) method of clustering on geographic areas because it gives larger standard errors than the replicate weights, likely because clustering on geography is too conservative a measure when analyzing a variable such as interstate migration that is not highly correlated across neighboring households. When we combine estimates for multiple years, we calculate separate point estimates for each year, take the unweighted average across years, and calculate the standard error assuming the estimates in different years are independent. (Because the CPS is a rotating panel of addresses, this assumption is not strictly correct, but the available sampling information in the public-use files does not allow us to easily relax it.)

2. Pre-2005 changes in American Community Survey procedures

This section describes how changes in procedures in the American Community Survey (ACS) before 2005 may have affected the estimated migration rate.

The ACS initially mails a survey form to sample housing units, then tries to telephone those who do not return the mailed form, and finally sends field representatives to personally interview a subsample of those who are not reached by mail or phone (U.S. Census Bureau, 2009, chap. 7). Estimated migration rates are lower for mail and telephone respondents than for in-person respondents, probably because the mail and telephone surveys are less likely to reach recent movers (Koerber, 2007). (Koerber, 2007, finds that migration rates vary little by response mode in the CPS because the CPS does not use mail surveys and is more likely than the ACS to have up-to-date phone numbers for respondents.) Thus, any variation in the percentage of respondents interviewed in person could translate into variation in the estimated migration rate.

From 2000, when the ACS became a large national demonstration project, through 2005, the first year of full implementation, survey procedures changed at least four times in

ways that could have affected the rate of in-person interviews and hence the migration rate. First, in 2002, budget constraints caused the ACS to conduct no surveys in July and to skip telephone and in-person follow-ups in June (Garrett and Williams, 2006). Second, in mid-2002, the rate of telephone interviews increased because the ACS obtained more telephone numbers from the decennial census and was able to contact more housing units by phone (Garrett and Williams, 2006). Third, in 2004, budget constraints caused the ACS to skip telephone and in-person follow-ups in January (U.S. Census Bureau, 2012). Fourth, through 2004, all housing units that were not reached by telephone or mail had a one-in-three chance of inclusion in the personal-interview subsample (U.S. Census Bureau, 2004), but since 2005, the probability of a personal interview has varied by census tract and by whether the housing unit has an address where mail can be delivered (U.S. Census Bureau, 2005, 2009).

The pre-2005 public-use data do not indicate the response mode, and the data since 2005 do not distinguish between telephone and in-person interviews, so we cannot adjust the data to account for these effects.

3. Details on patterns of migration

The main paper shows that many composition effects do not explain the decrease in interstate migration over the 1991–2011 period. This appendix examines these composition effects in more detail.

A. One-digit industry and occupation categories used to standardize population in figure 5

We use the following one-digit industry categories in figure 5 of the main paper:

1. Unknown
2. Agriculture/forestry/fishing
3. Mining/construction
4. Durable manufacturing
5. Nondurable manufacturing
6. Transportation/utilities
7. Trade
8. Finance/insurance/real estate
9. Services
10. Government

11. Not in labor force

We use the following one-digit occupation categories in figure 5 of the main paper:

1. Executive, administrative, and managerial
2. Professional specialty
3. Technicians and related support
4. Sales
5. Administrative support
6. Service
7. Farming, forestry, and fishing
8. Precision production, craft, and repair
9. Operators, fabricators, and laborers
10. Military
11. Unemployed not classified
12. Not in labor force

B. Demographics

Figure 1(a) shows that the decline in migration is virtually the same when we adjust for many different compositional changes by holding population distributions of various variables fixed at the 1991 distribution.

Figures 1(b), 1(c), and 1(d) show that migration rates have fallen at all education levels, for people of all marital statuses, and for both single-earner and multiple-earner households. Thus, although the demographics of the U.S. population have changed in many ways since 1991, these changes have no power for explaining the decline in interstate migration. In particular, the findings on marital status and number of earners demonstrate that the fall in migration is not due to changes in the number of “tied stayers” (Gemici, 2011; Guler, Guvenen, and Violante, 2012; Mincer, 1978) who cannot move because their partners cannot move.

C. Occupation and industry effects

Over the past several decades, the service sector has expanded while manufacturing has declined. If workers’ mobility rates differed across industries, this sectoral shift could produce a decline in migration. However, figure 2(a) shows that service-industry workers

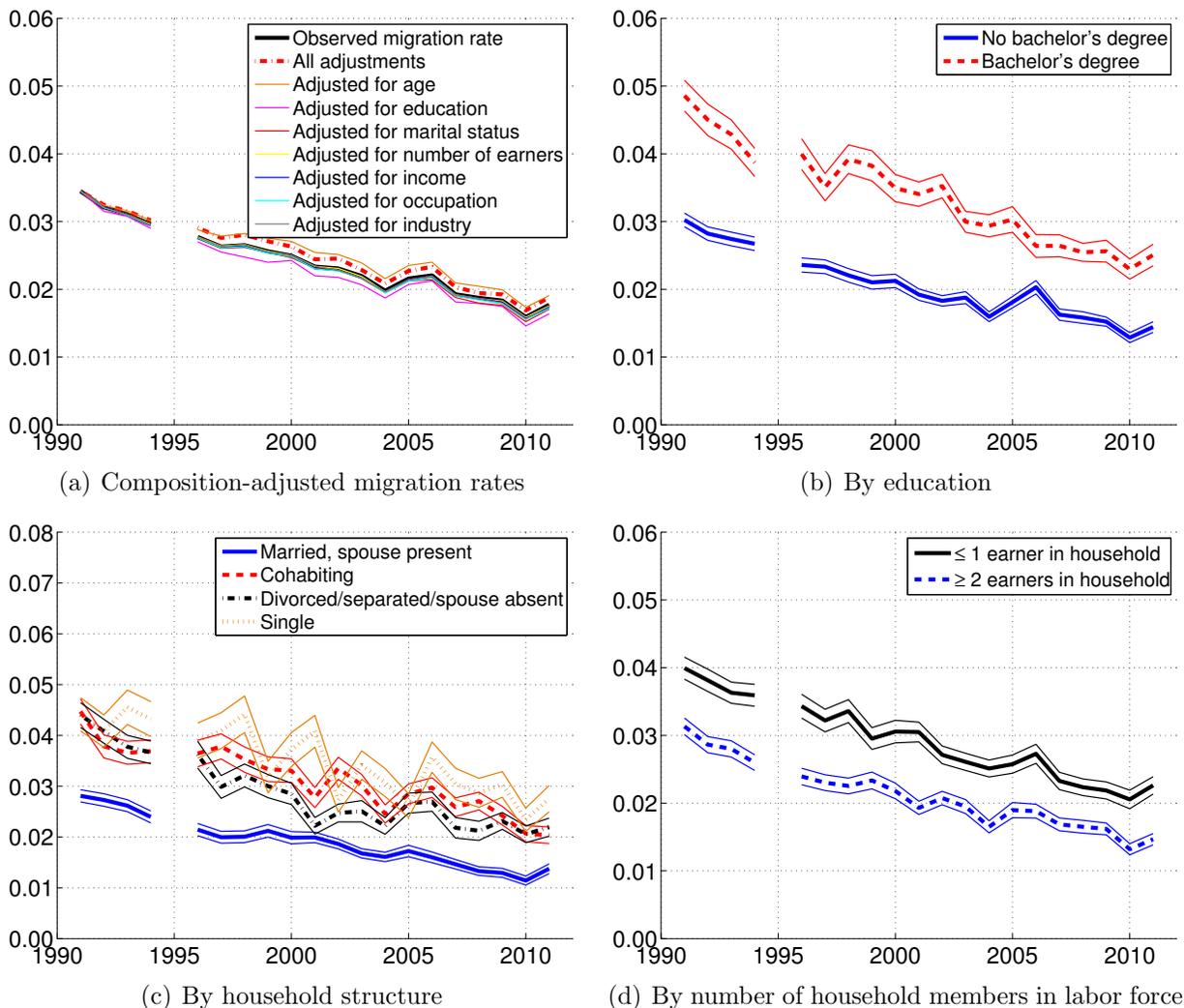


Figure 1: Time series of interstate migration by population subgroups.

Source: Authors' calculations from CPS micro data. The sample is restricted to working-age adults. Composition-adjusted rates hold the following variables constant at their 1991 distribution: respondent's age (single years), respondent's education (single years), respondent's marital status (four categories shown in figure 1(c)), number of labor force participants in respondent's household (two categories shown in figure 1(d)), and real income per capita of respondent's household (20 equal-population bins in 1991). The thin lines in figures 1(b), 1(c), and 1(d) show 1-standard-error confidence bands around point estimates.

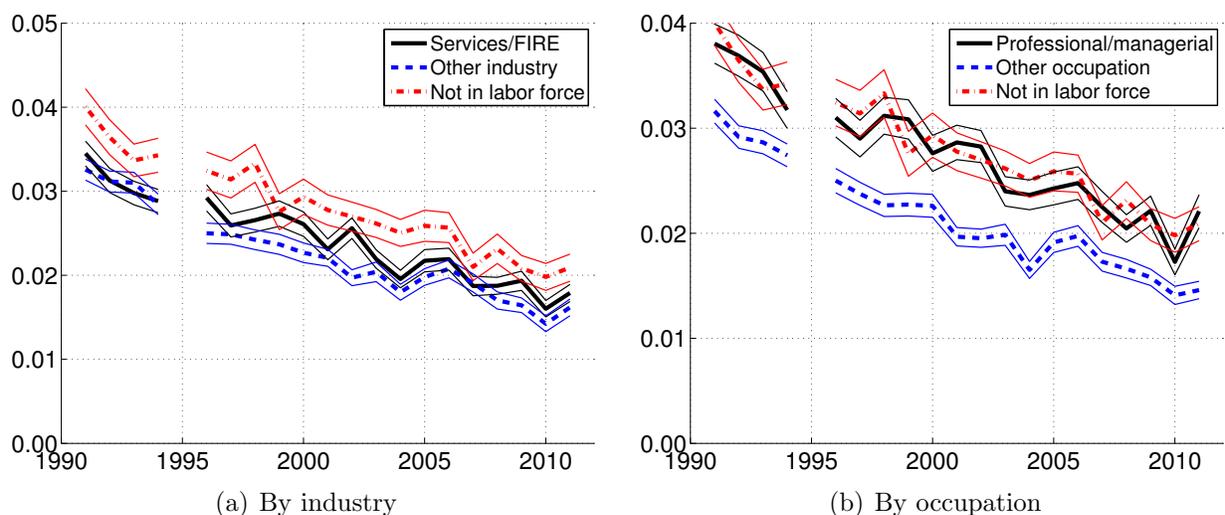


Figure 2: Migration rate by industry and occupation.

Source: Authors' calculations from CPS micro data. The sample is restricted to working-age adults. The thin lines show 1-standard-error confidence bands around point estimates. Professional/managerial occupations include all managerial and professional specialty occupations, and service/FIRE industries include all service industries and finance, insurance, and real estate as coded by King et al. (2010).

have approximately the same mobility as workers in other industries and that mobility has declined in parallel for workers in all industries. Further, figure 1(a) shows that when we hold the industry distribution fixed at the 1991 distribution, the migration trend does not change. Thus, the rise of the service sector seems to be an unlikely candidate for explaining the decline in migration.

Another hypothesis is that new communications technologies reduce migration by allowing some workers to do their jobs from anywhere in the country, instead of having to live in the city where their employer has its operations. These changes affect some occupations much more than others. But figure 2(b) shows that the migration rate for professional and managerial workers — who may have the most opportunities to work remotely — has declined only slightly more than the migration rate for workers in other occupations. Thus, we must seek an explanation for decreased migration that applies to all workers, not just those who can do their jobs over the Internet.

D. Income effects

The recent recession notwithstanding, the United States has grown wealthier since 1991. If living in one place for a long time is a normal good, the rise in incomes could cause

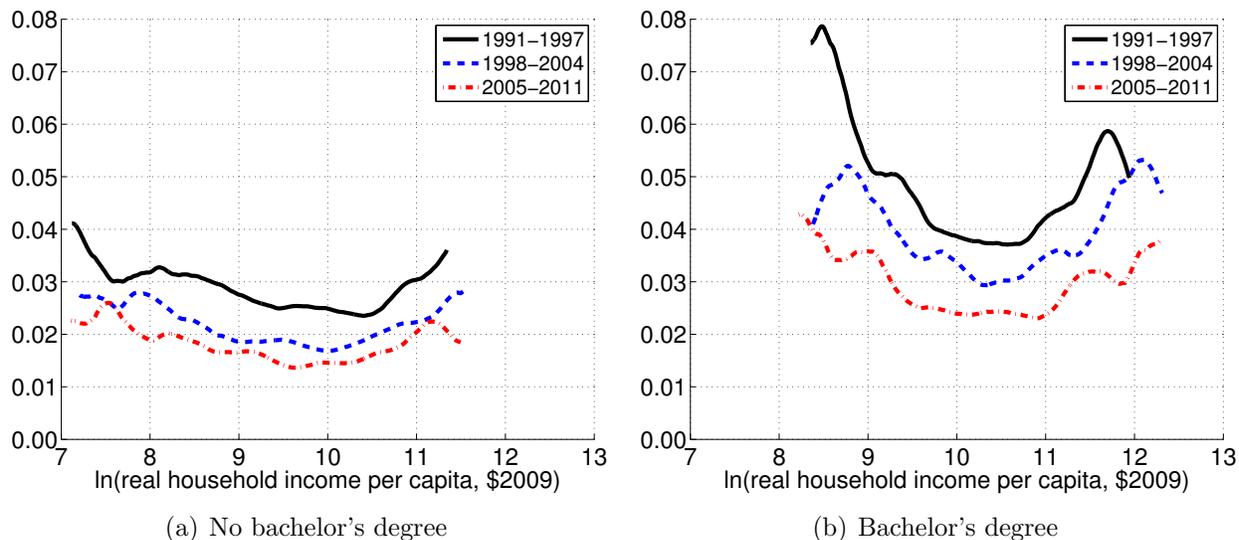


Figure 3: Migration rate by real household income, controlling for age.

Source: Authors' calculations from CPS micro data. The sample is restricted to working-age adults. The lines are local linear regression estimates of the migration rate as a function of income, controlling for age indicator variables in a partially linear model using the method of Yatchew (1997). Estimates use bandwidth 0.15 and Epanechnikov kernel. The graph is truncated at the 1st and 99th percentiles of the income distribution, conditional on education, in each period.

a fall in migration. Figure 3 tests this hypothesis by estimating the migration rate as a function of real household income per capita, controlling for age. We obtain the graph by estimating a partially linear model in which migration depends linearly on a full set of age indicator variables and nonparametrically on income: $migration = \mathbf{x}\boldsymbol{\beta} + f(income) + \epsilon$, where \mathbf{x} represents the age indicators. We estimate the linear part using the method of Yatchew (1997) with tenth-order differencing, then estimate the nonparametric part with local linear regression. We normalize the age effects to have mean zero.

Controlling for age is important because young people tend to have lower incomes and migrate more. Even after we control for age, migration is indeed higher at the low end of the income distribution. However, migration also ticks up at the high end of the distribution, so if income gains were concentrated among the already well off, the overall rise in incomes would not necessarily reduce migration. In addition, the figure shows that migration rates fell uniformly across the income distribution, and figure 1(a) shows that holding the real income distribution constant would not change the overall migration rate. Thus, rising real incomes do not explain the fall in migration.

E. Regional effects

Throughout U.S. history, high migration rates have been associated with large flows from one part of the country to another, most prominently in the Great Migration of African Americans out of the South. Is the recent decline in migration merely the result of a change in flows into or out of one part of the country? The net migration rate shown in figure 1 of the main paper suggests not: Even if all net interstate migration were eliminated, gross flows would barely change. Figure 4 examines this question another way by disaggregating the gross migration rate by region. Figure 4(a) shows each region’s gross inflow rate: the fraction of people in the region who lived in a different state (whether in the region or outside it) one year ago. Similarly, figure 4(b) shows each region’s gross outflow rate: the fraction of people who lived in the region one year ago and have since moved to a different state (whether in the region or outside it). The graphs show that both immigration and outmigration have fallen substantially in all regions. Thus, the driving force in the decline in migration cannot be a simple change in Americans’ desire or ability to move to or from one particular part of the country.

Another possibility is that in some parts of the country, interstate migration is a poor proxy for migration between labor markets. If migration from cities to suburbs has fallen over time, and if interstate migration captures some urban-suburban moves, we could mistakenly conclude that moves between labor markets have fallen when in fact they have not. We conjecture that this problem is likely to be most severe in the New York metropolitan area, which extends to large parts of New Jersey and Connecticut, and the Washington, D.C., metropolitan area, which extends into Maryland and Virginia. However, figure 4(c) shows that the decline in migration is actually larger when we exclude all respondents who live in New York, New Jersey, and Connecticut in the survey year, and is virtually identical when we exclude respondents in Maryland, Virginia, and Washington, D.C. (We exclude the entirety of the states because the boundaries of the metropolitan areas have changed over time.)

4. MSA-occupation interactions in income

Figure 5 show trends in the cross-sectional variance $\hat{\sigma}_{\xi,t}^2$ of location-occupation interactions in earnings when we define geographic locations by metropolitan statistical areas (MSAs) rather than states.

Using MSAs instead of states poses challenges because some people do not live in

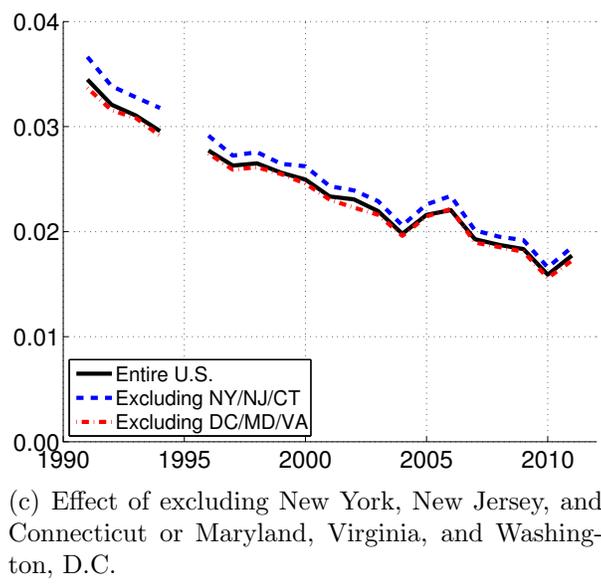
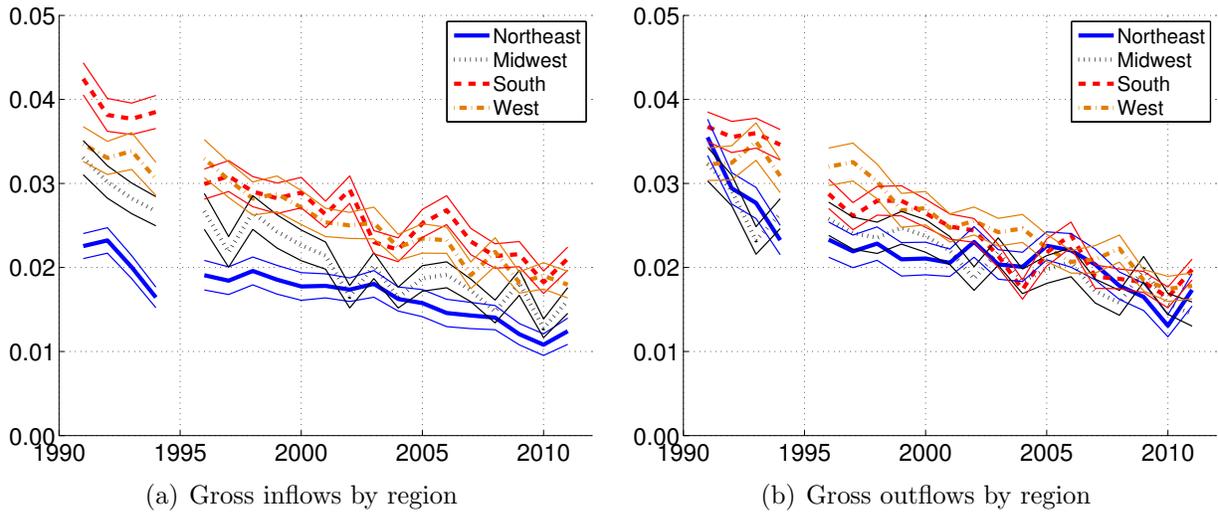


Figure 4: Regional patterns in migration.
 Source: Authors' calculations from CPS micro data. The sample is restricted to working-age adults. The thin lines in figures 4(a) and 4(b) show 1-standard-error confidence bands around point estimates.

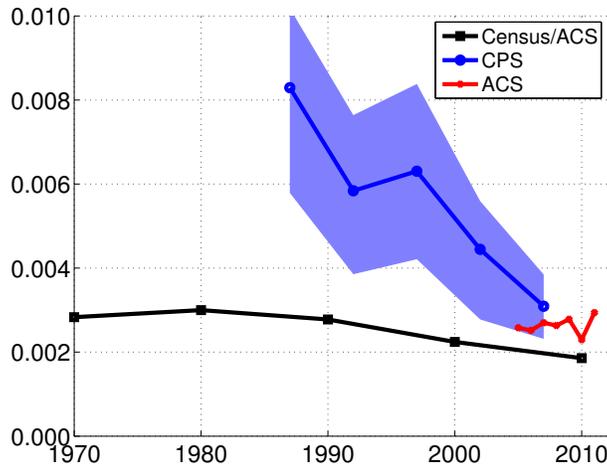


Figure 5: Cross-sectional variance of cell mean incomes across MSAs.

Source: Authors' calculations from Current Population Survey (CPS) micro data, 1977–2011; American Community Survey (ACS) micro data, 2005–2011; and decennial census micro data (1970, 1980, 1990, and 2000). Thick lines are bootstrap bias-corrected point estimates; shaded areas show bootstrap bias-corrected 90% confidence intervals around CPS point estimates. CPS estimates are averaged over five-year periods. Sample restricted to employed civilians ages 16 and over.

MSAs, not all MSAs are identified in public-use datasets, and MSA boundaries change over time. Also, MSAs can be identified in the CPS only starting in 1986. In the figure, we limit the sample for each dataset (census, CPS, or ACS) to respondents who live in MSAs that are identified in every year for that dataset. We do not adjust for changes in the counties that each MSA includes because the public-use data do not always show the respondent's county and because expansions of MSAs often reflect expansions of the geographic area that forms a single labor market. The MSA analysis uses the 1 percent form 1 and form 2 metro samples instead of the state samples from the 1970 census and drops the 1 percent sample from the 2000 census because it identifies relatively few MSAs.

Results with MSAs are similar to those using states, but the decline is now steeper in the CPS and less steep in the decennial census.

5. Alternative estimation of location-occupation interactions in income

This section describes alternative methods to estimate the variance of location-occupation interactions in equations (1) and (2) in the main paper. First, we give technical details on the parametric random-effects model reported in the paper. Then, we show that our results are robust to relaxing the random-effects orthogonality assumption.

A. Parametric random-effects model

To estimate the random-effects model parametrically, we assume ξ_{ost} is normally distributed with mean 0 and variance $\sigma_{\xi,t}^2$ and assume ϵ_{iost} is normally distributed with mean zero and constant variance. The large samples and large number of parameters make maximum likelihood estimators converge very slowly, so we estimate the posterior distribution of the parameters by Markov chain Monte Carlo using algorithm 16 of Chib (2001).

We use standard uninformative priors for the fixed effects a_{st} and b_{ot} and the coefficients β_t and standard weakly informative priors for the variances of ξ_{ost} and ϵ_{iost} . The prior has a larger effect on the posterior when the sample is smaller. We choose a prior for $\sigma_{\xi,t}^2$ centered on 0.001, which is smaller than any of the posterior means. Thus, the prior could produce a downward trend in the posterior estimates of $\sigma_{\xi,t}^2$ only if the samples were getting smaller over time; in reality, the sample sizes increase over time, which means that our results using this method are, if anything, an underestimate of the decrease in income dispersion.

For these estimates only, we do not use survey weights because we cannot derive the weighted likelihood without detailed information on sampling and weighting procedures (see, e.g., Gelman, 2007). The demographic controls \mathbf{x}_{iost} should absorb most of the subpopulation heterogeneity that survey weights are intended to account for.

B. Fixed effects instead of random effects

The main paper treats location-occupation interactions as random effects that are orthogonal to observed worker characteristics and therefore estimates the interactions from the means within occupation-location-year cells of residuals from a Mincer regression. We now investigate whether relaxing the orthogonality assumption changes the results. In place of equations (1) and (2), we estimate

$$\ln y_{iost} = a_{st} + b_{ot} + \mathbf{x}'_{iost} \beta_t + \xi_{ost} + \epsilon_{iost}, \quad (1')$$

where, as before, y_{iost} is the wage, salary, and self-employment income of worker i in occupation o , state or MSA s , and year t ; a_{st} is a location-year fixed effect; b_{ot} is an occupation-year fixed effect; \mathbf{x}_{iost} is a vector of controls, including sex, dummy variables for single year of education, and a quartic polynomial in potential experience; and ϵ_{iost} is independently and identically distributed across workers with mean zero. However, in this model, ξ_{ost} is an

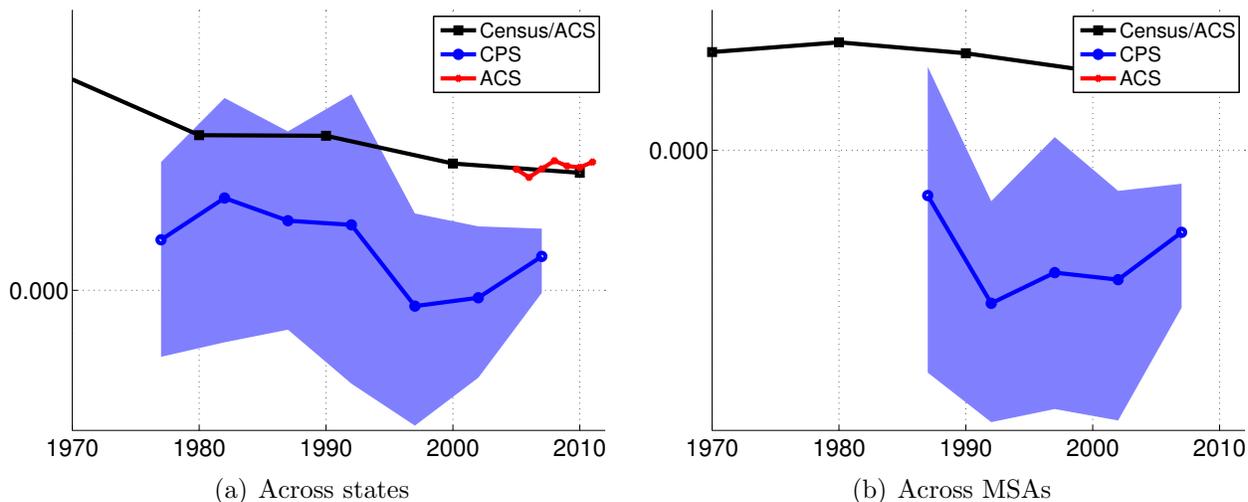


Figure 6: Cross-sectional variance of cell mean incomes.

Source: Authors' calculations from CPS micro data, 1977–2011; ACS micro data, 2005–2011; and decennial census micro data (1970, 1980, 1990, and 2000). Thick lines are bootstrap bias-corrected point estimates; shaded areas show bootstrap bias-corrected 90 percent confidence intervals around CPS point estimates. CPS estimates are averaged over five-year periods. The sample is restricted to employed civilians ages 16 and over.

occupation-location-year dummy variable rather than a random effect. Because we now allow for arbitrary correlation between ξ_{ost} and \mathbf{x}_{iost} , instead of forcing those correlations to be zero as in the main paper, this alternative approach potentially reduces the precision of our estimates.

Figure 6 shows our estimates of the cross-sectional variance of ξ_{ost} using the alternative estimation method. The trends in the census and ACS are quite similar to those in the main paper. The CPS data still show a downward trend, as in the main paper, but the confidence intervals become much wider, as expected.

6. Additional measures of dispersion of quantities of workers

This section reports the Theil segregation index for segregation of industries across states and of both industries and occupations across MSAs.

Figure 7 shows the index for segregation of industries across states. As with the occupation segregation measures in the main paper, states' distributions of workers across industries have become more similar. In results not shown here, we found that the decline is not solely due to the shift from manufacturing to services; an index of segregation of detailed industries within manufacturing also falls over the past four decades.

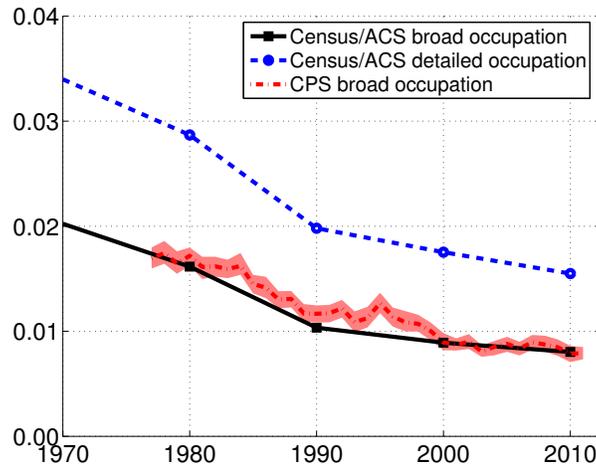


Figure 7: Theil indices of segregation of industries across states.

Source: Authors' calculations from CPS micro data, 1981–2011; ACS micro data, combined 2006–2010 sample; and decennial census micro data (1970, 1980, 1990, and 2000). Detailed industry categories are three-digit codes for the IND1990 variable in Ruggles et al. (2010). Broad industry categories are one-digit codes listed in subsection 3.A. The sample is restricted to employed civilians ages 16 and over and is weighted by the number of workers. Point estimates are bootstrap bias corrected. Shaded areas show bootstrap bias-corrected 90 percent confidence intervals around CPS point estimates. The 90 percent confidence intervals for census and ACS estimates are too small to be visible.

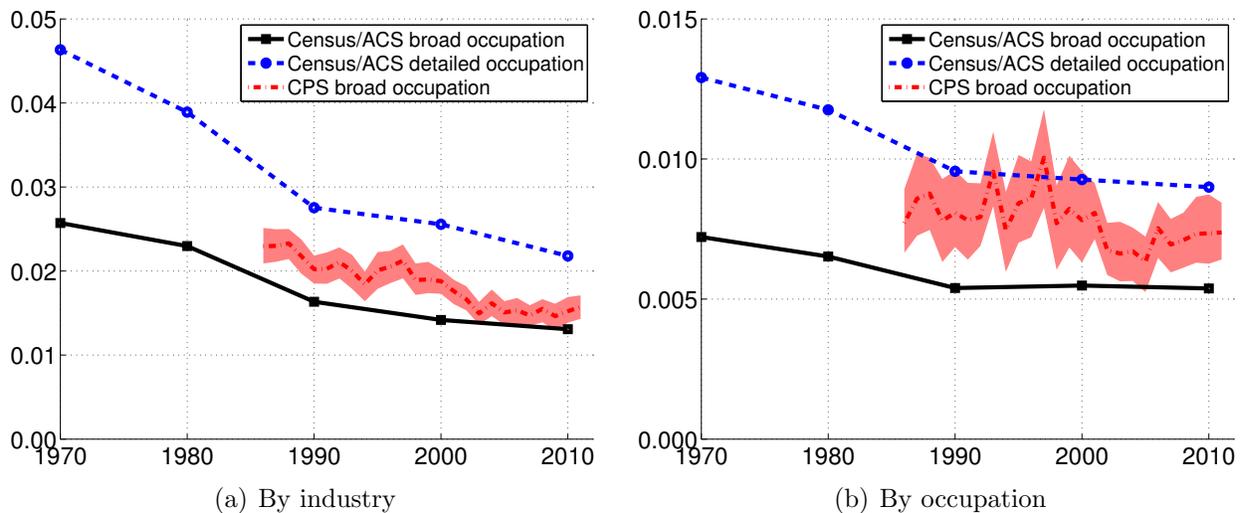


Figure 8: Theil indices of segregation of industries and occupations across MSAs.

Source: Authors' calculations from CPS micro data, 1981–2011; ACS micro data, combined 2006–2010 sample; and decennial census micro data (1970, 1980, 1990, and 2000). Detailed industry and occupation categories are three-digit codes for the IND1990 and OCC1990 variables in Ruggles et al. (2010). Broad industry and occupation categories are one-digit codes listed in subsection 3.A. The sample is restricted to employed civilians ages 16 and over and is weighted by the number of workers. Figures include only MSAs that can be identified in every year for a given dataset. Point estimates are bootstrap bias corrected. Shaded areas show bootstrap bias-corrected 90 percent confidence intervals around CPS point estimates. The 90 percent confidence intervals for census and ACS estimates are too small to be visible.

Figure 8 shows the Theil indices for occupational and industrial segregation across metropolitan statistical areas. As with states, the distributions of workers across industries and occupations in MSAs have become more similar over time.

7. Details on SIPP and census data on repeat and return migration

SIPP. We use data from the 1985, 1986, 1987, 1988, 1990, 1991, 1992, 1993, 1996, 2001, 2004, and 2008 Survey of Income and Program Participation (SIPP) panels. We omit the 1984 panel because it uses different groupings of small states that would cause us to lose information and the 1989 panel because it was followed for less than two full years.

The SIPP interviews respondents every four months. Thus, the wave 1 interview tells us the respondent’s location at baseline, the wave 4 interview tells us the location in year $t + 1$, and the wave 7 interview tells us the location in year $t + 2$. We restrict the sample to respondents who (a) had data collected in all of waves 1, 4, and 7 and (b) as of wave 4 were working-age adults (as defined in our CPS sample) and lived in households where no one was a member of the military. Restriction (a) implies that we exclude anyone who moves and is not followed by the SIPP. Restriction (b) makes our sample as comparable as possible to the working-age adults sample from the CPS (for which we measure migration between $t - 1$ and t among people who are working-age adults in nonmilitary households at t).

We construct a combined state code for respondents in Maine and Vermont, and a second combined state code for respondents in Alaska, Idaho, Iowa, Montana, North Dakota, South Dakota, and Wyoming, because states within these groups cannot be distinguished in some years in the public-use data files.

We weight the data by the SIPP panel weights. For panels where the Census Bureau constructed several panel weights, we use the panel weight that corresponds to the first two years of the survey. For pre-2001 panels, we use the half-sample and stratum codes provided with the public-use files to account for the survey sampling design when we estimate confidence intervals. Starting with the 2001 panel, we use the panel replicate weights to estimate confidence intervals.

Census. We use the 1 percent form 1 and form 2 state samples from the 1970 census, the 5 percent samples from the 1980 and 1990 censuses, and the 5 percent and 1 percent samples from the 2000 census, all obtained from Ruggles et al. (2010). We restrict the sample to

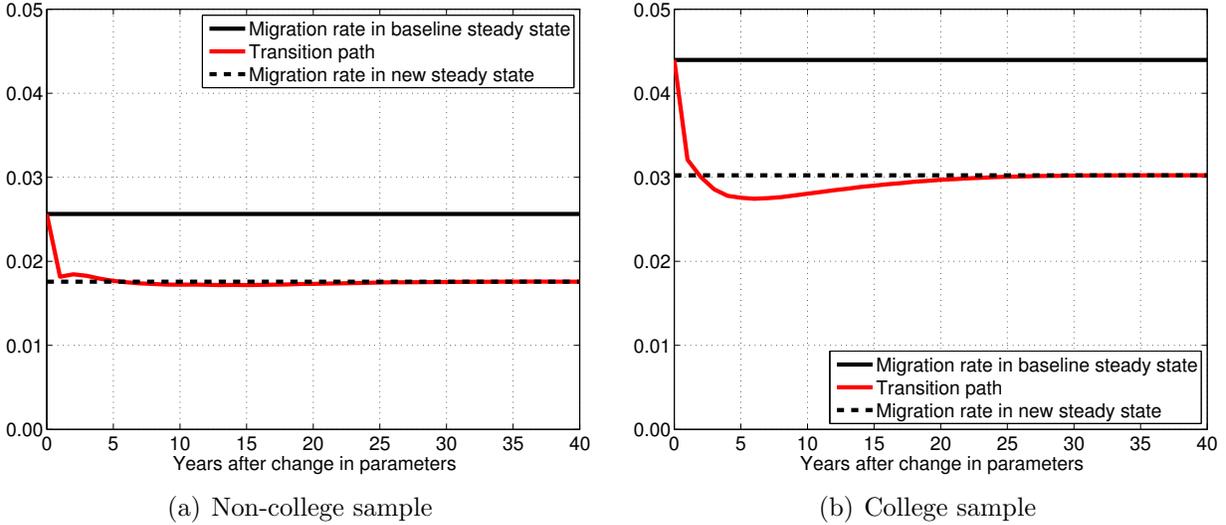


Figure 9: Transition dynamics of the average migration rate after a change in geographic specificity and information.

respondents who were born in the United States, lived in the United States five years earlier, do not have imputed birth states or locations five years earlier, and are working-age adults in nonmilitary households at the time of the census. We use the stratum and cluster codes provided by Ruggles et al. (2010) to account for the census public-use sample design when we estimate confidence intervals.

8. Transition dynamics

In the main paper, we claimed that it was appropriate to measure θ in the same years as the migration rate and to compare steady states with different parameters because convergence to the steady state is rapid in our model. In this appendix, we explore the model’s transition dynamics to provide evidence for this claim.

Figure 9 shows how the average migration rate in the model evolves after we change the geographic specificity of occupations and the amount of information agents have. In the experiment, agents begin in the steady state corresponding to our baseline calibration. Then we impose two one-time, unexpected, permanent parameter changes: a decrease in θ from 1.15 to 1.10 and an increase in the precision of every agent’s beliefs corresponding to six months of information.

As the figure shows, the migration rate reaches the level of the new steady state

within a few years. This rapid convergence occurs because the only potentially slow-moving endogenous state variable in our model is the precision of agents' beliefs, but our experiment directly shocks this precision.

Convergence in our model would be slower if we assumed that the new parameters applied only to cohorts newly reaching working age and not to agents who were of working age before the parameter change. Such an assumption would be quite strong, however. For example, it would mean assuming that workers who were 20 years old when the geographic specificity of occupations changed would ignore that change for the rest of their lives.

In any event, our model's overall ability to explain a large fraction of the decline in migration is not sensitive to what we assume about whether parameter changes affect new cohorts or all cohorts. This timing assumption has little effect on transition dynamics related to information, because changes in information mainly affect young agents, and after a few years, all young agents are from cohorts that reached working age after the parameter change. Thus, the timing assumption mainly has an impact on transition dynamics related to a change in θ , the geographic specificity of occupations. In additional experiments, we found that the migration rate reaches the new steady-state value after 20 to 30 years when the new parameters apply only to newborn cohorts. Therefore, if we assume new parameters affect only new cohorts, we should measure the change in θ over a period 20 to 30 years earlier than the period when we measure a decline in migration. Figure 6 of the main paper shows that the decline in the geographic specificity of occupations across states from 1970 to 1990, as measured in the census and ACS, was as large as the decline from 1990 to 2010.¹ As a result, our comparison of steady states based on 1991 and 2011 parameters gives quantitatively similar results to what we would obtain if we calibrated the model to the change in θ between 1970 and 1990, simulated transitions assuming that the new parameters affect only newborn cohorts, and compared the simulated migration rates in 1991 and 2011.

References

Chib, Siddhartha, 2001, "Markov Chain Monte Carlo Methods: Computation and Inference."
In *Handbook of Econometrics*, vol. 5, ed. James J. Heckman and Edward Leamer, 3569–

¹In addition, improvements in information have continued over many decades, and it seems reasonable to assume that the changes from 1970 to 1990 (related to airline and telephone deregulation) were similar in magnitude to those between 1991 and 2011. But such an assumption is not necessary because, as noted, the model converges rapidly after a change in information even if that change affects only new cohorts.

3649. Amsterdam: Elsevier.

Davern, Michael, Arthur Jones Jr., James Lepkowski, Gestur Davidson, and Lynn A. Blewett, 2006, “Unstable Inferences? An Examination of Complex Survey Sample Design Adjustments Using the Current Population Survey for Health Services Research.” *Inquiry* 43(3), 283–297.

Garrett, B. Dale, and Andre L. Williams, 2006, “Meeting 21st Century Demographic Data Needs — Implementing the American Community Survey: Report 6: The 2001–2002 Operational Feasibility Report of the American Community Survey.” Available at http://www.census.gov/acs/www/Downloads/library/2004/2004_Garrett_01.pdf.

Gelman, Andrew, 2007, “Struggles with Survey Weighting and Regression Modeling.” *Statistical Science* 22(2), 153–164.

Gemici, Ahu, 2011, “Family Migration and Labor Market Outcomes.” Manuscript, New York University.

Guler, Bulent, Fatih Guvenen, and Giovanni L. Violante, 2012, “Joint-Search Theory: New Opportunities and New Frictions.” *Journal of Monetary Economics* 59(4), 352–369.

King, Miriam, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick, 2010, Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota.

Koerber, Kin, 2007, “Comparison of ACS and ASEC Data on Geographic Mobility: 2004.” Manuscript, U.S. Census Bureau. Available at http://www.census.gov/acs/www/library/by_year/2007/.

Mincer, Jacob, 1978, “Family Migration Decisions.” *Journal of Political Economy* 86(5), 749–773.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek, 2010, Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota.

- U.S. Census Bureau, 2004, “Accuracy of the Data (2004).” Available at http://www.census.gov/acs/www/Downloads/data_documentation/Accuracy/accuracy2004.pdf.
- U.S. Census Bureau, 2005, “Accuracy of the Data (2005).” Available at http://www.census.gov/acs/www/Downloads/data_documentation/Accuracy/accuracy2005.pdf.
- U.S. Census Bureau, 2009, “Design and Methodology: American Community Survey.” Available at http://www.census.gov/acs/www/Downloads/survey_methodology/acs_design_methodology.pdf.
- U.S. Census Bureau, 2012, American Community Survey, “Response Rates — Data,” December 17, accessed Jan. 17, 2013. Available at https://www.census.gov/acs/www/methodology/response_rates_data/.
- Yatchew, A., 1997, “An Elementary Estimator of the Partial Linear Model.” *Economics Letters* 57(2), 135–143.