



A sharp drop in interstate migration? Not really

New data procedures led to misperception of dramatic decline in U.S. population mobility

April 2011

Greg Kaplan
University of Pennsylvania

Sam Schulhofer-Wohl
Federal Reserve Bank of Minneapolis

ABSTRACT

Many policymakers have expressed concern that unemployment remains high, in part, because the once highly mobile American worker has suddenly become unable or unwilling to move across the country for a job. This paper shows that this concern is unnecessary: Contrary to popular belief, interstate migration did not fall substantially during the Great Recession; in fact, interstate migration has probably been overestimated in the past.

The misperception of a sharp drop in migration is due to a statistical artifact. In 2006, the Census Bureau changed the methods for handling data for people who do not answer migration questions in the Bureau's Current Population Survey. We find that this change in data-handling procedures—not any change in actual migration patterns—explains nearly half of the reported decrease in interstate migration between 2000 and 2010. Many factors are undoubtedly to blame for high unemployment in the United States, but a sharp drop in migration is not among them.

A sharp drop in interstate migration? Not really

New data procedures led to misperception of dramatic decline in U.S. population mobility¹

Greg Kaplan and Sam Schulhofer-Wohl

April 2011

Introduction

The news reports have been grim:

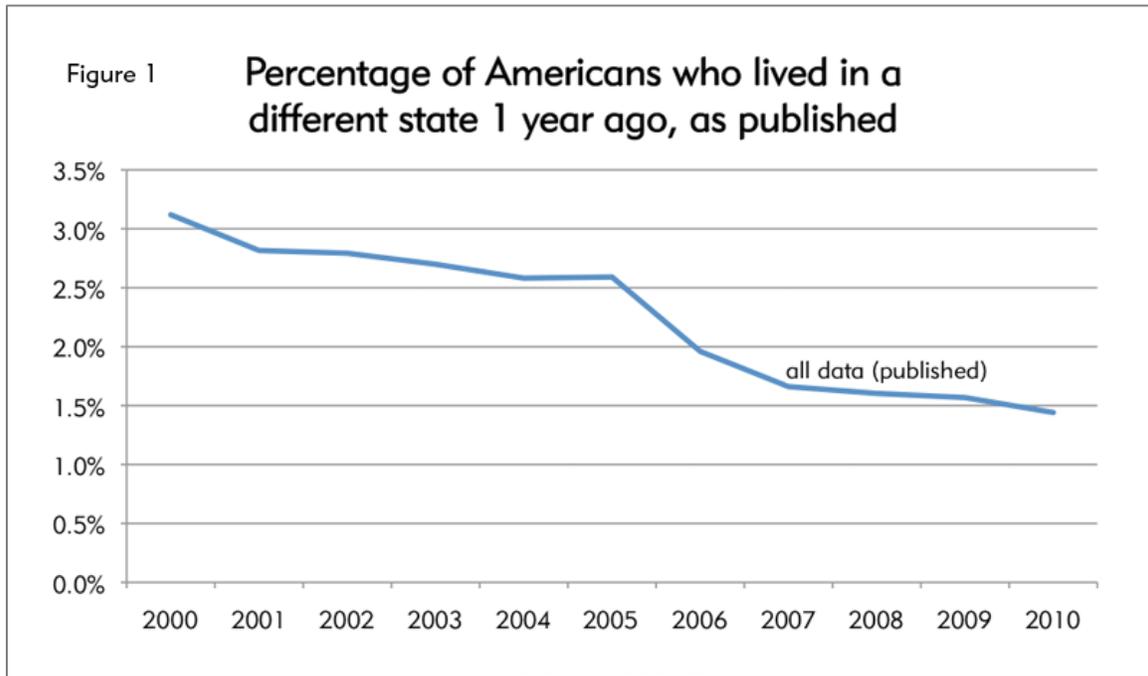
- “Slump Creates Lack of Mobility for Americans” (*New York Times*, April 2009)²
- “The recession is claiming yet another victim: Americans’ near-constitutional right to pick up and move to a better job.” (*Washington Post*, July 2010)³
- “One of the hallmarks of the American worker has been mobility—the speed with which people ... have moved to find opportunities. But the recession of the last two years has produced a profound change, creating conditions that have tethered many people where they are.” (*Los Angeles Times*, December 2009)⁴

Low migration has raised worries not just in the media, but among policymakers. The fundamental fear is that if unemployed workers don’t move to states with stronger job markets, the economy could remain stalled for years. Citing such arguments, the International Monetary Fund recently blamed “slower inter-state migration, likely related to the housing crash,” for the persistent high level of U.S. unemployment.⁵ Leaders of the U.S. Treasury and the Federal Reserve System have expressed similar concerns.

A closer look at the data reveals that such concerns are unfounded. Our research shows that what one demographer called the “Great American Migration Slowdown”⁶ never really happened.

Figure 1 shows the data that got reporters, researchers and policymakers so worried: the annual interstate migration rate for the past decade, as calculated by the U.S. Census Bureau and published on its Web site.⁷ According to this graph, the migration rate apparently plummeted in 2006 from a relatively high plateau earlier in the decade.

Figure 1: Percentage of Americans who lived in a different state one year ago, as calculated and published by the Census Bureau, 2000 to 2010



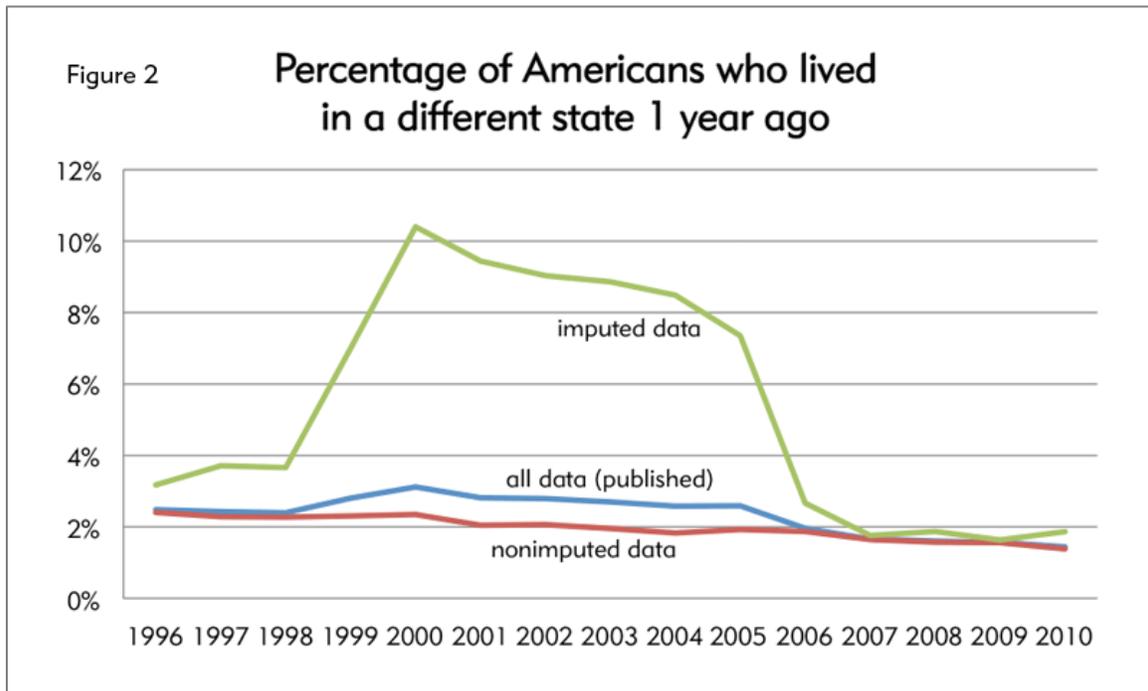
But this graph is misleading. The data for it come from the Census Bureau’s Current Population Survey (CPS), but a 2006 change in how the bureau analyzed the data lowered the measured rate of migration. Indeed, it could be argued that the new method is more accurate than the old and that previous bureau reports overestimated interstate migration rates while understating local migration. In any case, our research shows that interstate migration rates have remained on a slow downward trend over the past decade, and *there has been no dramatic change in this trend.*

How did a change in data analysis cause such misperceptions? About 10 percent of respondents in the CPS don’t answer the bureau’s questions about where they lived a year ago; to calculate the migration rate, the bureau has to guess whether these people migrated and, if so, from where. Statisticians call these guesses “imputations.” In 2006, the bureau changed the way it calculates the imputations. This change in methods—*not* any actual change in migration patterns—turns out to be responsible for much of the recent decline in reported migration rates. The change explains 90 percent of the reported decrease in interstate migration between 2005 and 2006, and 42 percent of the decrease between 2000 (the recent high-water mark) and 2010.

Figure 2 illustrates the problem. The figure shows the interstate migration rate for *all* CPS respondents and, separately, the rate for those with original, *nonimputed* data and the rate for those with imputed data. From 1996 to 1998 and from 2006 to 2010, the rate for respondents with imputed data is only slightly higher than the rate for respondents with

original data, and the rate for all respondents is likewise very close to the rate for respondents with original data. But from 1999 to 2005, the interstate migration rate for respondents with imputed data is *three to five times* the rate for respondents with nonimputed data.

Figure 2: Percentage of Americans who lived in a different state one year ago—all data, nonimputed data and imputed data, 1996-2010.



In this paper, we explain why statisticians impute answers for people who don't answer survey questions, describe how the Census Bureau's imputation methods changed, document how the change affected estimates of migration rates and discuss some policy implications of our results. This paper is an informal discussion, written in a question-and-answer format. Readers interested in technical details should refer to our Federal Reserve Bank of Minneapolis Working Paper 681, "Interstate Migration Has Fallen Less Than You Think: Consequences of Hot Deck Imputation in the Current Population Survey." The views expressed here are those of the authors and not necessarily of others in the Federal Reserve System.

Why does the Census Bureau "invent" answers to survey questions?

In 2010, of every 100 people surveyed for the CPS, 87 said they hadn't moved between states, 1 said they had moved between states, and 12 didn't answer questions about migration. What we assume about those 12 people out of every hundred matters a lot: If they all migrated, the migration rate would be much higher than if none of them

migrated. We could ignore the 12 entirely—and, in fact, we suggest later in this paper that ignoring the nonrespondents turns out to be a good solution in this case.

But scholars, including ourselves, usually think it's a bad idea to ignore people who don't answer survey questions. The reason is that nonrespondents are often very different from respondents. For example, young workers move more often than other people, and young workers might also be too busy with their careers to diligently answer every question on the survey. In that case, the people who do answer the migration questions won't be representative of all Americans—they will be older and less mobile than the average person. To get an accurate picture of migration, it's important to take account of these differences. The Census Bureau does so by “inventing” (or imputing) what it believes to be accurate responses for those questions that weren't answered.

If the CPS has so many skipped questions, why not use a different data set?

The CPS migration data are a unique and invaluable resource for research on internal migration in the United States. CPS migration data have been published annually since 1948, the longest-running migration data series the nation has. Unlike other large data sets, such as the decennial census and the American Community Survey, the CPS can precisely measure fluctuations in migration from one year to the next. The CPS also allows researchers to study the relationship between migration and a vast number of other characteristics of workers and households. In addition, the CPS provides a representative sample of the entire U.S. population, unlike data from the Internal Revenue Service (which cover only income-tax filers) or moving companies (which cover only people wealthy enough to hire movers).

Other data sets do support our argument that migration has not fallen drastically in recent years. Migration rates in the nonimputed CPS data—but not in the imputed data—are consistent with rates calculated from other data sets.

How does the Census Bureau impute answers for people who skip questions?

The Census Bureau uses a method called “hot deck allocation.” Participants' survey forms are fed through a computer one at a time. If someone didn't answer a particular question, the computer looks back through the forms it processed previously until it finds the most recently processed survey participant who's similar to the nonrespondent but who *did* answer the question. The previously processed forms are called the hot deck.

Let's say Mr. Smith, a 33-year-old black homeowner living in the Midwest, didn't answer a particular question. The computer might look through the hot deck for the most recently processed 33-year-old black homeowner in the Midwest who did answer—Mr. Jones, say—and then copy Mr. Jones' answer onto Mr. Smith's form. The person whose

missing answer is filled in is known as a recipient; the person whose answer is used is known as a donor. The process is called imputation through hot deck allocation.

Why does hot deck allocation help?

Basically, hot deck allocation guesses that people who have similar demographic characteristics, such as age, race and region of residence, are also similar in other characteristics, such as migration. It's not a perfect method, of course, but it's much better than assuming nonrespondents are similar to respondents, which is rarely true. For example, hot deck allocation ensures fairly accurate results even if young people are more likely to skip the migration questions.

What changed in 2006?

The Census Bureau changed the order in which it feeds the survey forms through the computer. That sounds innocuous, but it wasn't. If the forms go into the computer in a different order, then the computer picks different forms from the hot deck when it needs to fill in missing answers. The new computer processing order reduced the reported interstate migration rate by imputing fewer interstate moves and more local moves to nonrespondents.

How exactly does the processing order matter?

For migration data, the imputation procedure fills in missing answers to two questions: Did the person live in the same home one year ago? If not, where did the person live one year ago? Once these variables are filled, additional variables are calculated that categorize movers as having moved within a county, between counties in the same state, between states or from abroad, based on the distance between the respondent's current location and his or her (possibly imputed) location one year ago.

Since 2006, the Census Bureau has processed CPS surveys in geographic order. So if someone skips the migration questions, the computer usually copies an answer from a donor who lives nearby. Since long-distance migration is rare, the donor's location one year ago is also usually close to the recipient's current location. Thus, if the computer imputes that the recipient moved, it usually imputes a local move.

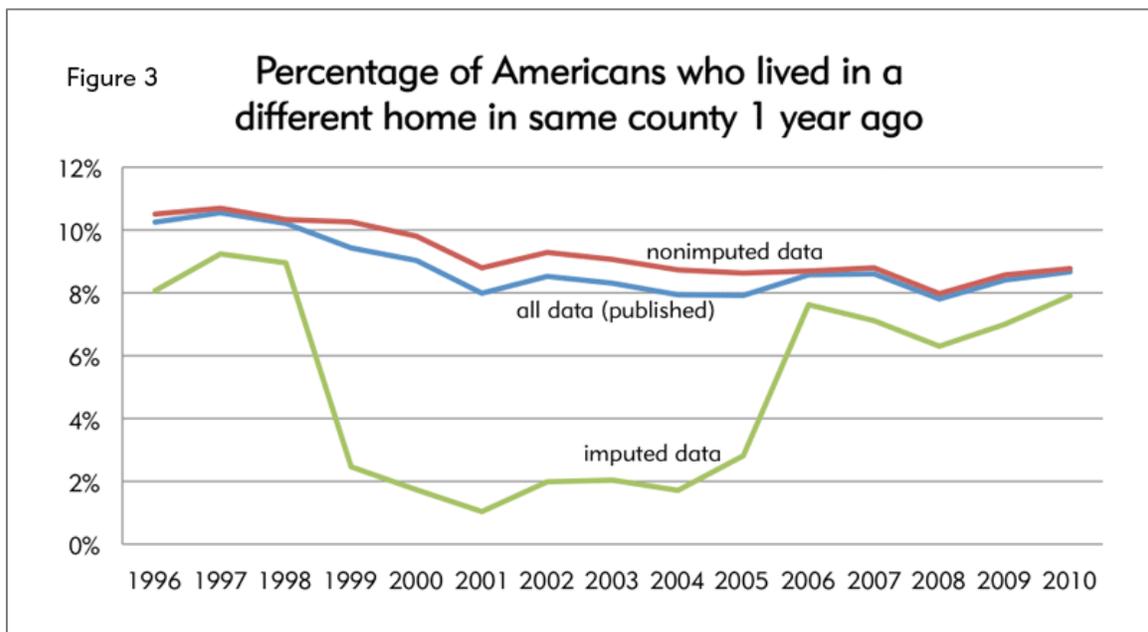
But *before* 2006, the data were not exactly in geographic order.⁸ That meant donors lived farther from recipients, donors' locations one year ago were also farther from recipients' current locations and recipients were more likely to have imputed interstate moves.

For example, suppose a person in Minneapolis fails to answer the migration questions and is matched with a donor who moved, so that the nonrespondent is coded as a mover. If imputations are done in geographic order, the donor will probably also come from Minneapolis, and the donor’s location one year ago was also probably near Minneapolis (since long-distance migration is statistically infrequent). The geographic procedure will thus usually impute that the nonrespondent made a *local* move, not an interstate move.

However, if imputations are not done in exact geographic order, the donor is more likely to live far away from Minneapolis—in Sioux Falls, say. Because interstate migration is quite unusual, the donor’s location a year ago was probably near Sioux Falls. But the pre-2006 procedure will inaccurately guess that the nonrespondent in Minneapolis made an *interstate* move, from South Dakota to Minnesota, not because the nonrespondent is similar to a respondent who moved a long distance (from Sioux Falls to Minneapolis), but rather because the nonrespondent is similar to a respondent who moved a short distance in South Dakota.

In Figure 2, we show how the pre-2006 procedure imputed many more interstate moves than the new procedure. It’s also interesting to look at local moves. Figure 3 shows that from 1999 to 2005, the within-county migration rate was depressed among respondents with imputed data. Thus, the pre-2006 imputation procedure spuriously imputed long-distance moves when they were more likely local moves. Our research has found that the change in imputation procedures had little effect on the total migration rate, because the decrease in imputed interstate moves in 2006 canceled out an increase in imputed within-county moves.

Figure 3: Percentage of Americans who lived in a different home in the same county one year ago—all data, nonimputed data and imputed data, 1996-2010



Which migration rate estimates are more accurate: pre-2006 or post-2006?

We agree with the Census Bureau that the procedure in use since 2006 is likely to produce the most reliable estimates of migration rates. The old procedure likely overstated the rate of interstate migration and understated the rate of local migration.

Do the problems with imputed migration data mean it's always a bad idea to use imputed data?

Absolutely not! As we discuss above, imputing missing data is better than dropping nonrespondents in almost all circumstances, because nonrespondents might differ from respondents, and those differences should not be ignored. In addition, we believe the current statistics including imputations are the most accurate estimates of interstate migration—more accurate than either the nonimputed data or the estimates using the old imputation procedure. The problem here is not with imputation itself; it's just that a *change* in imputation methods created a misleading trend in the migration statistics. In sum, imputation is extremely useful, but researchers must analyze the data carefully when imputation procedures change from one year to the next.

How has interstate migration changed in recent years?

The change in imputation procedures in 2006 means that simple comparisons of pre-2006 and post-2006 Census Bureau data do not accurately measure trends in interstate migration. Fortunately, there is a simple way to make accurate comparisons: By ignoring imputed data we can eliminate the problem introduced when the Census Bureau switched methods mid-way through the 2000s.

Since 2006, the interstate migration rate including imputed data has been virtually identical to the rate using only nonimputed data. Figure 4 illustrates this point by reproducing Figure 2 without the imputed-data migration rate, to show only the contrast between overall and nonimputed rates. Because the interstate migration rates using nonimputed data and using all data have been virtually identical for the past five years, we think that the rate using nonimputed data is a reliable guide before 2006 as well. We can study trends in the overall interstate migration rate by focusing on the rate in nonimputed data, and doing so removes the fluctuations induced by changes in the imputation procedure.

We have also checked whether the nonimputed data are a reliable guide to the trend over time by constructing our own imputations for every year from 1996 to 2010, using a method that does not change over time. When we do this, we find that the estimated migration rate using our imputation method is virtually identical to the rate in nonimputed data.

Figure 4: Percentage of Americans who lived in a different state one year ago—all data, nonimputed data, 1996-2010.

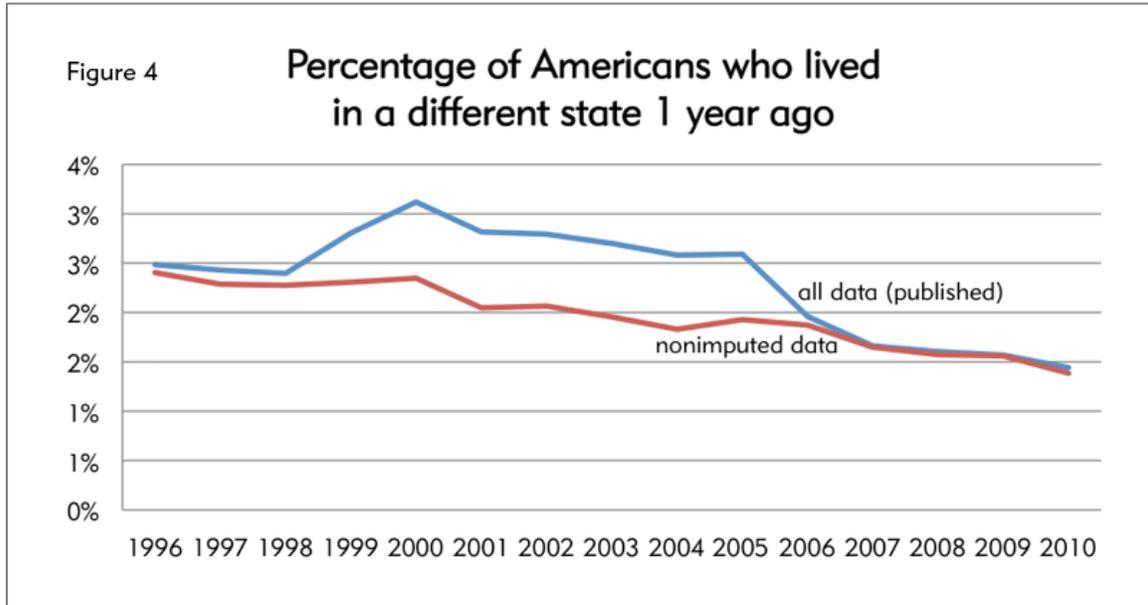


Figure 4 shows that, once we remove the effect of changes in the imputation procedure, the interstate migration rate has hewed closely to a smooth downward trend for the past 15 years. *With* imputations included, the rate peaked at 3.12 percent in the 2000 survey, fell to 2.59 percent by the 2005 survey, plummeted to 1.96 percent in the 2006 survey, and is now down to 1.44 percent. *Without* imputations, the rate was 2.35 percent in the 2000 survey, 1.93 percent in the 2005 survey, 1.87 percent in the 2006 survey, and 1.38 percent in the 2010 survey. Thus, the change in imputation procedures explains nine-tenths of the 0.63 percentage point drop from 2005 to 2006 and four-tenths of the 1.68 percentage point drop from 2000 to 2010—the apparent declines in mobility that have been of such concern to policymakers and others.

Did interstate migration fall during the Great Recession?

As Figure 4 shows, interstate migration has been trending downward for many years. But, relative to that trend, there was no additional decrease in interstate migration during the December 2007-June 2009 recession. To see this, it's important to know how the CPS data are collected. The survey is taken in February through April each year and asks people whether they moved in the previous 12 months. Thus, the data points corresponding to migration during the recession are the data points for the 2008, 2009 and 2010 surveys. Figure 4 shows that, in the nonimputed data—that is, the data we consider to be most accurate—migration fell faster than trend in the 2007 survey year, well *before* the recession began in December.

Will the long-run decline in migration cause problems for the labor market?

This is an important focus for ongoing research. Many factors might affect whether people migrate. It can be costly to sell a house and buy a new one. Some people have a strong preference for living near family and friends. Others may want to move to a better climate or a place where there are more jobs. Which of these factors are mainly responsible for the downward trend in migration will determine whether we should expect falling migration rates to hurt the labor market.

On the one hand, if migration is falling because it is becoming more difficult to sell a house, then it may be difficult for unemployed workers to move to where the jobs are. On the other hand, in the past, the United States has usually experienced high migration rates when some regions of the country had much stronger economies than others. Historically, people migrated from the South to manufacturing centers in the North, and from rural farming communities to big cities. Regional economic disparities have been declining in the United States for some time. If that's why migration is falling, there is little cause for concern: Migration is lower than in the past because it's unneeded, not because it's costly.

How should the government respond if low migration is hurting the labor market?

Determining the best policies requires a careful analysis of why migration has fallen—and, as a first step, accurate data on how much migration actually *has* fallen. If research shows that migration actually has fallen significantly, that low migration is hurting labor markets *and* that, for example, the high cost of selling a home is an important factor inhibiting migration, it still isn't necessarily the case that the government can or should intervene. Studying and designing potential policies is another important area for research. Our hope is that the corrected data series on migration that we provide in our working paper will help make such research possible.

¹ This policy paper is based on Kaplan, Greg, and Sam Schulhofer-Wohl, 2011, “Interstate Migration Has Fallen Less Than You Think: Consequences of Hot Deck Imputation in the Current Population Survey,” Federal Reserve Bank of Minneapolis Working Paper 681.

² Roberts, Sam, 2009, “Slump Creates Lack of Mobility for Americans,” *New York Times*, April 23, p. A1.

³ Fletcher, Michael A., 2010, “Few in U.S. move for new jobs, fueling fear the economy might get stuck, too,” *Washington Post*, July 30, p. A1.

⁴ Lee, Don, 2009, “Need work; won’t travel: Americans’ mobility has dropped to its WWII level, reducing employers’ options,” *Los Angeles Times*, Dec. 25, p. A17.

⁵ Batini, Nicoletta, Oya Celasun, Thomas Dowling, Marcello Esteveão, Geoffrey Keim, Martin Sommer, and Evridiki Tsounta, 2010, “United States: Selected Issues Paper,” International Monetary Fund Country Report 10/248.

⁶ Frey, William H., 2009, “The Great American Migration Slowdown: Regional and Metropolitan Dimensions,” Brookings Institution report.

⁷ “Geographical Mobility/Migration,” U.S. Census Bureau, accessed Oct. 28, 2010, <http://www.census.gov/population/www/socdemo/migrate.html>.

⁸ By “not exactly,” we mean that while the order of processing was geographic, the Census Bureau sorted surveys geographically *within* particular subsamples of the data, rather than sorting the entire data set. (The portion of the CPS that contains migration data, the Annual Social and Economic Supplement, consists of several subsamples.)