Web Appendix for "The (Un)importance of Geographical Mobility in the Great Recession"

Siddharth Kothari, Itay Saporta-Eksten and Edison Yu

A Methodology

In this section we provide a full derivation of our stock-flow equations followed by details on how we compound the counterfactual over time. We then formally describe the homeownership correction conducted in Section 7. Finally, we discuss the conditions which are required to get an upper bound for the effect of mobility on unemployment.

A.1 Stock-Flow Equation

The main equation that relates the unemployment stock of workers to the flows into and out of unemployment is given below[1]

$$\begin{aligned} U_t &= U_{t-1} \\ &- (U_{H,t-1} \cap MJ_{t-1} \cap E_t) \\ &- \left(U_{H,t-1} \cap \overline{MJ}_{t-1} \cap E_t\right) \\ &- (U_{R,t-1} \cap E_t) \\ &+ Other\ Net\ inflow_{t-1}. \end{aligned} \tag{A.1}$$

Here U_{t-1} is the set of unemployed at time $t-1$; $U_{H,t-1}$ is the set of unemployed homeowners at time $t-1$; $U_{R,t-1}$ is the set of unemployed renters at time $t-1$; MJ_{t-1} is the set of people who move for jobs between time $t-1$ and t (people whose labor market outcomes improve as a result of moving); $\overline{MJ}_{t-1}$ is the complement of the set MJ_{t-1} and includes all non-movers and people who move but their moves were not job related; $Other\ Net\ inflow_{t-1}$ refers to all separations between employer and worker which lead to a transition of the worker from employment to unemployment between $t-1$ and t, and to net inflows to unemployment from out of the labor force; and E_t is the set of employed people at time t.

Using shares and reorganizing, equation (A.1) becomes

$$\begin{aligned} U_t &= U_{t-1} \\ &- U_{H,t-1} m_{t-1}^{UH} j_{t-1}^{UH} \left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) \\ &- U_{H,t-1} \bar{e}_{t-1}^{UH} \\ &- U_{R,t-1} e_{t-1}^{UR} \\ &+ E_{t-1} \delta_{t-1} \end{aligned} \tag{A.2}$$

Where, $m_{t-1}^{UH} = \frac{U_{H,t-1} \cap M_{t-1}}{U_{H,t-1}}$ – moving rate among unemployed homeowners (M_{t-1} are all movers between t-1 and t);

$j_{t-1}^{UH} = \frac{U_{H,t-1} \cap MJ_{t-1}}{U_{H,t-1} \cap M_{t-1}}$ – job-related moves as a share of all moves among unemployed homeowners;

[1] Because we only care about the level of unemployment and not about tracking specific individuals from unemployment to employment, it would have been more accurate to replace each set in equation (A.1) by its cardinality (e.g., replace $U_{H,t-1} \cap MJ_{t-1} \cap E_t$ by $|U_{H,t-1} \cap MJ_{t-1} \cap E_t|$). However, to simplify notation, we omit the cardinality notation throughout.

$e^{UH}_{t-1} = \frac{U_{H,t-1} \cap MJ_{t-1} \cap E_t}{U_{H,t-1} \cap MJ_{t-1}}$ – share of unemployed homeowners who move for jobs and become employed by the next period;

$\bar{e}^{UH}_{t-1} = \frac{U_{H,t-1} \cap \overline{MJ}_{t-1} \cap E_t}{U_{H,t-1} \cap \overline{MJ}_{t-1}}$ – share of other unemployed homeowners who become employed by the next period

$e^{UR}_{t-1} = \frac{U_{R,t-1} \cap E_t}{U_{R,t-1}}$ – share of unemployed renters who become employed by the next period; and

$\delta_{t-1} = \frac{E_{t-1} \cap U_t}{E_{t-1}}$ – share of employed who become unemployed by the next period.[2]

Collapsing the last three terms and naming it N_t, equation (A.2) yields equation (1) in the paper.

$$U_t = U_{t-1} - U_{H,t-1} m^{UH}_{t-1} j^{UH}_{t-1} \left(e^{UH}_{t-1} - \bar{e}^{UH}_{t-1} \right) + N_t \tag{A.3}$$

where, $N_t = -U_{H,t-1} \bar{e}^{UH}_{t-1} - U_{R,t-1} e^{UR}_{t-1} + E_{t-1} \delta_{t-1}$

A.2 Counterfactual Exercise

As noted in the paper, the counterfactual exercise is conducted by first recovering $\hat{N}_t$ and then boosting the mobility rate of homeowners from m^{UH}_{t-1} to $m^{UH,C}_{t-1}$.[3] For the first time period in which we run the counterfactual, the counterfactual unemployment level is given by

$$U^C_1 = U_0 - U_{H,0} m^{UH,C}_0 j^{UH}_0 \left(e^{UH}_0 - \bar{e}^{UH}_0 \right) + \hat{N}_1. \tag{A.4}$$

The subscript 0 represents the base period and 1 the first period of the counterfactual. For subsequent periods, the counterfactual unemployment rate is calculated by compounding the counterfactual from the previous period:

$$U^C_t = U^C_{t-1} - U^C_{H,t-1} m^{UH,C}_{t-1} j^{UH}_{t-1} \left(e^{UH}_{t-1} - \bar{e}^{UH}_{t-1} \right) + \hat{N}_t. \tag{A.5}$$

Note that we also need to keep track of the counterfactual level of unemployed homeowners ($U^C_{H,t-1}$) because it appears on the right-hand side of equation (A.5). This is done using an equation that is similar to equation (A.3), but for unemployed homeowners. We can write the stock flow equation for $U_{H,t}$ as in the following:

$$U_{H,t} = U_{H,t-1} - U_{H,t-1} m^{UH}_{t-1} j^{UH}_{t-1} \left(e^{UH}_{t-1} - \bar{e}^{UH}_{t-1} \right) + NN_t.$$

As with $\hat{N}_t$, we can back out the residual term NN_t by using data for the first two terms:

$$N\hat{N}_t = U_{H,t} - U_{H,t-1} + U_{H,t-1} m^{UH}_{t-1} j^{UH}_{t-1} \left(e^{UH}_{t-1} - \bar{e}^{UH}_{t-1} \right).$$

Then we can boost the probability of move to get the counterfactual unemployment level within the group of homeowners as

$$U^C_{H,t} = U^C_{H,t-1} - U^C_{H,t-1} m^{UH,C}_{t-1} j^{UH}_{t-1} \left(e^{UH}_{t-1} - \bar{e}^{UH}_{t-1} \right) + N\hat{N}_t, \tag{A.6}$$

where for the base period $U^C_{H,0} = U_{H,0}$.

A.3 Homeownership Correction

In the CPS, we do not observe homeownership status prior to the move but only post-move. We therefore approximated the share of homeowners who moved, $H_{t-12} \cap M_{t-12}$, by using $H_t \cap M_{t-12}$. In this section we show how to construct a correction for the change in homeownership (conducted in Section 7 of the paper). This correction allows us to proxy $H_{t-12} \cap M_{t-12}$ better at the cost of focusing on the population of 16- to 65-year-olds instead of only individuals in the labor force.

[2] To ease notation, we assume that the labor force is fixed, and $Other\ Net\ inflow_{t-1}$ therefore includes only transitions from employment to unemployment.

[3] In some of the exercises we also take the counterfactual of j^{UH}_t. All the derivations are identical, and can be thought of as setting $j^{UH}_0 = 1$ and redefinig m^{UH}_{t-1} as job-related mobility and $m^{UH,C}_{t-1}$ as counterfactual job-related mobility.

We are interested in recovering the number of homeowners at $t-12$ who moved between $t-12$ and t. This number is given by

$$H_{t-12} \cap M_{t-12} = H_{t-12} \cap M_{t-12} \cap H_t + H_{t-12} \cap M_{t-12} \cap R_t \quad (A.7)$$

while the proxy measure which we have used for our baseline calculations is

$$H_t \cap M_{t-12} = H_{t-12} \cap M_{t-12} \cap H_t + R_{t-12} \cap M_{t-12} \cap H_t$$

To get an accurate measure of $(H_{t-12} \cap M_{t-12})$, we need to subtract from our baseline measure the number of renters who moved and became homeowners $(R_{t-12} \cap M_{t-12} \cap H_t)$ and add the number of homeowners who moved and became renters $(H_{t-12} \cap M_{t-12} \cap R_t)$. Our correction is based on the observation that transitions from renting to owning status and vice verse would most likely be accompanied by a move. Formally we assume

$$H_{t-12} \cap M_{t-12} \cap R_t \approx H_{t-12} \cap R_t \quad (A.8)$$
$$R_{t-12} \cap M_{t-12} \cap H_t \approx R_{t-12} \cap H_t \quad (A.9)$$

This allows us to relate our measures of interest ($R_{t-12} \cap M_{t-12} \cap H_t$ and $H_{t-12} \cap M_{t-12} \cap R_t$) to the change in the numbers of homeowners and renters in the population given by

$$\begin{aligned} \Delta H_t &= H_t - H_{t-12} = R_{t-12} \cap H_t - H_{t-12} \cap R_t + NH_t \\ \Delta R_t &= R_t - R_{t-1} = H_{t-12} \cap R_t - R_{t-12} \cap H_t + NR_t \end{aligned}$$

where NH_t and NR_t are net new homeowners and net new renters (new in the sense of new to the population we are looking at), respectively. We can use assumptions (A.8) and (A.9) (second equality in equation (A.10)) to write equation (A.7) as

$$H_{t-12} \cap M_{t-12} \approx H_t \cap M_{t-12} - \Delta H_t + NH_t = H_t \cap M_{t-12} + \Delta R_t - NR_t \quad (A.10)$$

While $H_t \cap M_{t-12}$, ΔH_t and ΔR_t are observed, NH_t and NR_t are not observed. To measure NH_t, we would need to know both how many people in the labor force today were not in the labor force in the last period and also their homeownership status. Since the CPS does not track the labor force status of individuals in the previous year, we cannot measure NH_t when working with the labor force as our population. However, if we redefine our population as all people in the age group of 16- to 65-year-olds, then we can measure NH_t and NR_t

$$\begin{aligned} NH_t &= H_t \cap [age_t = 16] - H_{t-12} \cap [age_{t-12} = 65] \\ NR_t &= R_t \cap [age_t = 16] - R_{t-12} \cap [age_{t-12} = 65] \end{aligned}$$

Given NH_t and NR_t, all the right-hand side variables in equation (A.10) are observed, implying that we can calculate $H_{t-12} \cap M_{t-12}$.

A.4 Upper Bound - Propositions and Proof

In this section, we formalize under precisely what assumptions the results in the paper are an upper bound for the effect of mobility on unemployment. In particular, we show that there are two factors that make our counterfactual unemployment rate *smaller* than what the true counterfactual unemployment rate should be (i.e., we obtain an *upper bound* for the effect of mobility on unemployment).

First, our counterfactual exercise ignores the effect of changing mobility on net flows which are *not* directly related to job-related mobility (N_t). When we conduct the counterfactual exercise, we keep N_t constant and ignore the fact that higher mobility should also impact these net flows into unemployment. We start by showing under which assumptions, when compounding over more than one period, the measured N_t underestimates the flows into unemployment. This implies that the counterfactual unemployment calculated by our exercise is lower than the counterfactual unemployment when the relation between mobility and N_t is taken into account. This result, which we refer to as *methodological upper bound,* is formalized in Proposition 1.

Second, when conducting the counterfactual exercise, we use values for j_{t-1}^{UH}, and $\left(e_{t-1}^{UH}-\bar{e}_{t-1}^{UH}\right)$ that are higher than their values in the population. Moreover, for a given choice of counterfactual mobility, we try to boost the difference $\left(m_{t-1}^{UH,C}-m_{t-1}^{UH}\right)$, which further exaggerates the effect of mobility on unemployment. We refer to this result as *measurement upper bound* and formalize it in Proposition 2.

A.4.1 Methodological Upper Bound

In the paper, we defined a job finding rate for all unemployed homeowners who do not move for job and called it $\bar{e}_{t-1}^{UH}$. Here we further split this group into two groups - first, the group of unemployed homeowners who move for jobs only in the counterfactual i.e. who move for job in the counterfactual but do not do so in the observed world; second, the group who do not move for job both in the observed and in the counterfactual world. We define the job finding rate of these two groups as $\bar{e}_{t-1}^{UH}$ and $e_{t-1}^{UH,NM}$ respectively. The reason for this change is that our counterfactual exercise changes the composition of the group of unemployed homeowners who do not move for jobs. In the counterfactual world, the size of the group shrinks as some of the unemployed in this group move for jobs and are assigned the higher job-finding rate e_{t-1}^{UH}. By defining two job finding rates, we clarify how this composition shift impacts our counterfactual exercise.

Proposition 1: *Define U_t^{CI} to be the true counterfactual that takes into account the effect of change in mobility rate on N_t. The following conditions are* sufficient *for U_t^C (as defined in equation (A.5)) to be a lower bound (upper bound for the effect of mobility), i.e. $U_t^C \leqslant U_t^{CI}, \forall t$:*

(C.1) δ_t, j_t^{UH}, and all finding rates (e^{UH}, $\bar{e}_t^{UH}$, $e_{t-1}^{UH,NM}$ and e_t^{UR}) are not changing in the counterfactual $\forall t$.

(C.2) $m_{t-1}^{UH,C}\, j_{t-1}^{UH}\left(e_{t-1}^{UH}-e_{t-1}^{UH,NM}\right)+e_{t-1}^{UH,NM}+\delta_{t-1}\leq 1$

Proof: We start by showing that $U_1^C = U_1^{CI}$ holds. Equation (A.2) becomes

$$\begin{aligned}
U_t = {} & U_{t-1} \\
& - U_{H,t-1} m_{t-1}^{UH} j_{t-1}^{UH}\left(e_{t-1}^{UH}-\bar{e}_{t-1}^{UH}\right) \\
& - U_{H,t-1} m_{t-1}^{UH,C} j_{t-1}^{UH}\bar{e}_{t-1}^{UH} \\
& - U_{H,t-1}\left[1-m_{t-1}^{UH,C} j_{t-1}^{UH}\right] e_{t-1}^{UH,NM} \\
& - U_{R,t-1} e_{t-1}^{UR} \\
& + E_{t-1}\delta_{t-1}
\end{aligned}$$

where for a given $m_{t-1}^{UH,C}$, the third line refers to the group of individuals who move for job in the counterfactual. Define

$$\hat{N}_t = U_t - U_{t-1} + U_{H,t-1} m_{t-1}^{UH} j_{t-1}^{UH}\left(e_{t-1}^{UH}-\bar{e}_{t-1}^{UH}\right)$$

and the counterfactual becomes

$$U_t^C = U_{t-1}^C - U_{H,t-1}^C m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH}-\bar{e}_{t-1}^{UH}\right)+\hat{N}_t$$

which for the first period implies

$$U_1^C = U_1 - U_{H,0}\left(m_0^{UH,C}-m_0^{UH}\right) j_0^{UH}\left(e_0^{UH}-\bar{e}_0^{UH}\right) \tag{A.11}$$

The true counterfactual for period 1, U_t^{CI} is given by

$$\begin{aligned}
U_1^{CI} = {} & U_0 \\
& - U_{H,0} m_0^{UH,C} j_0^{UH} e_0^{UH} \\
& - U_{H,0}\left[1-m_0^{UH,C} j_0^{UH}\right] e_0^{UH,NM} \\
& - U_{R,0} e_0^{UR} \\
& + E_0\delta_0
\end{aligned}$$

Adding and subtracting U_1 from this equation we get

$$U_1^{CI} = U_1 - U_{H,0}\left(m_0^{UH,C} - m_0^{UH}\right) j_0^{UH}\left(e_0^{UH} - \bar{e}_0^{UH}\right) = U_1^C \tag{A.12}$$

We show now how to obtain an upper bound with compounding. Define the differences of observed unemployment and (true) counterfactual unemployment as the following:

$$\begin{aligned}
\Delta U_t^C &\equiv U_t - U_t^C \\
\Delta U_{H,t}^C &\equiv U_{H,t} - U_{H,t}^C \\
\Delta U_t^{CI} &\equiv U_t - U_t^{CI} \\
\Delta U_{H,t}^{CI} &\equiv U_{H,t} - U_{H,t}^{CI}
\end{aligned}$$

Using this notations, we can express these differences as a set of first order linear difference equations

$$\begin{aligned}
\Delta U_t^C = {} & \Delta U_{t-1}^C + U_{H,t-1}\left(m_{t-1}^{UH,C} - m_{t-1}^{UH}\right) j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) \\
& -\Delta U_{H,t-1}^C m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right)
\end{aligned} \tag{A.13}$$

$$\begin{aligned}
\Delta U_{H,t}^C = {} & \Delta U_{t-1}^C + U_{H,t-1}\left(m_{t-1}^{UH,C} - m_{t-1}^{UH}\right) j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) \\
& -\Delta U_{H,t-1}^C m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right)
\end{aligned} \tag{A.14}$$

$$\begin{aligned}
\Delta U_t^{CI} = {} & \Delta U_{t-1}^{CI} + U_{H,t-1}\left(m_{t-1}^{UH,C} - m_{t-1}^{UH}\right) j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) \\
& -\Delta U_{H,t-1}^{CI} m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) - \Delta U_{H,t-1}^{CI} m_{t-1}^{UH,C} j_{t-1}^{UH} \bar{e}_{t-1}^{UH} \\
& -\Delta U_{H,t-1}^{CI}\left[1 - m_{t-1}^{UH,C} j_{t-1}^{UH}\right] e_{t-1}^{UH,NM} - \Delta U_{H,t-1}^{CI}\delta_{t-1}
\end{aligned} \tag{A.15}$$

$$\begin{aligned}
\Delta U_{H,t}^{CI} = {} & \Delta U_{H,t-1}^{CI} + U_{H,t-1}\left(m_{t-1}^{UH,C} - m_{t-1}^{UH}\right) j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) \\
& -\Delta U_{H,t-1}^{CI} m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) - \Delta U_{H,t-1}^{CI} m_{t-1}^{UH,C} j_{t-1}^{UH} \bar{e}_{t-1}^{UH} \\
& -\Delta U_{H,t-1}^{CI}\left[1 - m_{t-1}^{UH,C} j_{t-1}^{UH}\right] e_{t-1}^{UH,NM} - \Delta U_{H,t-1}^{CI}\delta_{t-1}
\end{aligned} \tag{A.16}$$

where, in the base period $\Delta U_0^C = \Delta U_{H,0}^C = \Delta U_0^{CI} = \Delta U_{H,0}^{CI} = 0$. It is easy to recognize that $\Delta U_t^C = \Delta U_{H,t}^C$, and $\Delta U_t^{CI} = \Delta U_{H,t}^{CI}$, $\forall t \geqslant 0$. Hence, if the ultimate goal is to sign $U_t^{CI} - U_t^C = \Delta U_t^C - \Delta U_t^{CI}$, it is sufficient to examine the sign of $\Delta U_{H,t}^C - \Delta U_{H,t}^{CI}$.

We prove the proposition by induction. We know that $\Delta U_{H,1}^C - \Delta U_{H,1}^{CI} = 0$.

Now assume that $\Delta U_{H,t-1}^C - \Delta U_{H,t-1}^{CI} \geqslant 0$ and $\Delta U_{H,t-1}^{CI} \geq 0$. If we can prove that $\Delta U_{H,t}^C - \Delta U_{H,t}^{CI} \geq 0$ and $\Delta U_{H,t}^{CI} \geq 0$, then we are done.

Taking the difference of equation (A.14) and (A.16), and rearranging gives

$$\begin{aligned}
\Delta U_t^C - \Delta U_t^{CI} = {} & \left(\Delta U_{H,t-1}^C - \Delta U_{H,t-1}^{CI}\right)\left[1 - m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right)\right] \\
& + \Delta U_{H,t-1}^{CI}\left[m_{t-1}^{UH,C} j_{t-1}^{UH} \bar{e}_{t-1}^{UH} + \left(1 - m_{t-1}^{UH,C} j_{t-1}^{UH}\right) e_{t-1}^{UH,NM} + \delta_{t-1}\right]
\end{aligned}$$

The induction assumptions were that $\Delta U_{H,t-1}^C - \Delta U_{H,t-1}^{CI} \geq 0$ and $\Delta U_{H,t-1}^{CI} \geq 0$. Therefore, all terms in the above equation are greater than zero. Therefore, $\Delta U_{H,t}^C - \Delta U_{H,t}^{CI} \geq 0$

To complete the proof we need to show that $\Delta U_{H,t}^{CI} \geq 0$. Rearranging equation (A.16) we get

$$\begin{aligned}
\Delta U_{H,t}^{CI} = {} & \Delta U_{H,t-1}^{CI}\left[1 - m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - e_{t-1}^{UH,NM}\right) - e_{t-1}^{UH,NM} - \delta_{t-1}\right] \\
& + U_{H,t-1}\left(m_{t-1}^{UH,C} - m_{t-1}^{UH}\right) j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right)
\end{aligned}$$

Notice that the above equation is simply an AR(1) in $\Delta U_{H,t}^{CI}$ but with coefficients which are time varying. In particular,

$$\Delta U_{H,t}^{CI} = b_{t-1}\Delta U_{H,t-1}^{CI} + a_{t-1}$$

where $b_{t-1} = 1 - m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - e_{t-1}^{UH,NM}\right) - e_{t-1}^{UH,NM} - \delta_{t-1}$ and $a_t = U_{H,t-1}\left(m_{t-1}^{UH,C} - m_{t-1}^{UH}\right) j_{t-1}^{UH}\left(e_{t-1}^{UH} - \bar{e}_{t-1}^{UH}\right) \geq 0$.

Under the induction assumption that $\Delta U_{H,t-1}^{CI} \geq 0$, a sufficient condition for $\Delta U_{H,t}^{CI} \geq 0$ is that $b_t \geq 0$ i.e. $m_{t-1}^{UH,C} j_{t-1}^{UH}\left(e_{t-1}^{UH} - e_{t-1}^{UH,NM}\right) + e_{t-1}^{UH,NM} + \delta_{t-1} \leq 1$.

We address assumption C.1 in footnote 10 in the paper. Note that this assumption gives sufficient, but not necessary conditions for $\hat{N}_t$ to overstate the ideal counterfactual residual flow into unemployment. Note that assumption (C.2) is a technical assumption related to compounding. In particular, it is required to rule out the pathological case in which increased mobility actually leads to higher unemployment (i.e., $U_t^{CI} > U_t$).

A.4.2 Measurement Upper Bound

We turn now to show that when we measure $\left(m_t^{UH,C} - m_t^{UH}\right)$, j_t^{UH} and $\left(e_t^{UH} - \bar{e}_t^{UH}\right)$ with an upward bias, we amplify the effect of mobility on unemployment even further. Proposition 2 shows this for a single period counterfactual. We find it useful to state the proposition for a single period because the bias for this case is very easy to calculate. The proposition and proof extending this result to multiple periods are available upon request.

Let $\tilde{m}_t^{UH}$, $\tilde{j}_t^{UH}$, $\tilde{\bar{e}}_t^{UH}$ and $\tilde{e}_t^{UH}$ be the measured values of m_t^{UH}, j_t^{UH}, $\bar{e}_t^{UH}$ and e_t^{UH} and $\tilde{m}_t^{UH,C}$ be the counterfactual mobility rate used. $\hat{N}_1$ and U_1^C are now given by

$$\hat{N}_1 = U_1 - U_0 + U_{H,0}\tilde{m}_0^{UH}\tilde{j}_0^{UH}\left(\tilde{e}_0^{UH} - \tilde{\bar{e}}_0^{UH}\right)$$

$$U_1^C = U_0^C - U_{H,0}\tilde{m}_0^{UH,C}\tilde{j}_0^{UH}\left(\tilde{e}_0^{UH} - \tilde{\bar{e}}_0^{UH}\right) + \hat{N}_1 \tag{A.17}$$

Proposition 2: *Let time* 0 *be the base period and* U_1^{CI} *be the true counterfactual that takes into account the effect of change in mobility rate on* N_1. *The following conditions are* sufficient *for* U_1^C *(as defined in equation (A.17)) to be a lower bound (upper bound for the effect of mobility), i.e.* $U_1^C \leqslant U_1^{CI}$*:*

(C.1) holds for $t = 0$

(C.3) $\tilde{m}_0^{UH,C} - \tilde{m}_0^{UH} \geq m_0^{UH,C} - m_0^{UH}$, $\tilde{j}_0^{UH} \geq j_0^{UH}$ and $\tilde{e}_0^{UH} - \tilde{\bar{e}}_0^{UH} \geq e_0^{UH} - \bar{e}_0^{UH}$

Proof: Equation (A.12) is unchanged, but equation (A.11) should be now written as

$$U_1^C = U_1 - U_{H,0}\tilde{j}_0^{UH}\left(\tilde{e}_0^{UH} - \tilde{\bar{e}}_t^{UH}\right)\left(\tilde{m}_0^{UH,C} - \tilde{m}_0^{UH}\right) \tag{A.18}$$

Taking the difference of equations (A.18) and (A.11) we get

$$\begin{aligned} U_1^{CI} - U_1^C &= U_{H,0}(\tilde{m}_0^{UH,C} - \tilde{m}_0^{UH})\tilde{j}_0^{UH}\left(\tilde{e}_0^{UH} - \tilde{\bar{e}}_t^{UH}\right) - U_{H,0}\left(m_0^{UH,C} - m_0^{UH}\right) j_0^{UH}\left(e_0^{UH} - \bar{e}_0^{UH}\right) \\ &= U_{H,0}\left[(\tilde{m}_0^{UH,C} - \tilde{m}_0^{UH})\tilde{j}_0^{UH}\left(\tilde{e}_0^{UH} - \tilde{\bar{e}}_t^{UH}\right) - \left(m_0^{UH,C} - m_0^{UH}\right) j_0^{UH}\left(e_0^{UH} - \bar{e}_0^{UH}\right)\right] \geq 0 \end{aligned}$$

where the last inequality is using assumption *(C.3)*.

The intuition for assumption (C.3) is as follows: $\tilde{m}_0^{UH,C} - \tilde{m}_0^{UH} \geq m_0^{UH,C} - m_0^{UH}$ and $\tilde{j}_0^{UH} \geq j_0^{UH}$ ensure that the size of the group of interest (movers for jobs) is overestimated. $\tilde{e}_0^{UH} - \tilde{\bar{e}}_0^{UH} \geq e_0^{UH} - \bar{e}_0^{UH}$ further ensures that the mobility-related flows out of unemployment for any given size of the group is also overestimated.

How large a bias is generated by our choice of parameters? We demonstrate it by showing the magnitude of the bias that our choice of $\tilde{e}_0^{UH} - \tilde{\bar{e}}_0^{UH}$ can generate. In our exercise we set this difference to 1, implying a monthly finding rate of 1 for movers for jobs and of 0 for all other homeowners. If all other parameters are set to their population values, our counterfactual exercise in the first period overstates the effect of mobility by

$$\frac{U_1 - U_1^C}{U_1 - U_1^{CI}} = \frac{1}{\left(e_0^{UH} - \bar{e}_0^{UH}\right)}$$

Keeping the finding rate for movers for jobs (e_0^{UH}) at the upper bound of 1, and assigning 0.23 (the average monthly hazard rate during the Great Recession calculated using the method in Shimer (2005)) to the finding rate for all homeowner ($\bar{e}_0^{UH}$), our counterfactual exercise overstates the effect of mobility on unemployment by about 30%.

Note that when choosing mobility rates, we boost both $\tilde{m}_0^{UH,C}$ *and* $\tilde{m}_0^{UH}$ proportionally (using the PSID to get to unemployed mobility from total mobility). This is an effort to make the difference $\tilde{m}_0^{UH,C} - \tilde{m}_0^{UH}$ also larger than the true population value.

B Data Appendix

We use data from three sources:

1. March CPS - We use the Annual Social and Economic Supplement (ASEC) to the Current Population Survey (CPS) as the main source of data for mobility and reason for move. The data from 1980 to 2011 comes from IPUMS-CPS.[4] For 2005 to 2012, we also use the NBER March CPS files to construct a panel of matched individuals over time and between monthly and March files, as matching is not possible using the IPUMS data.[5]

2. Monthly CPS Data - We use the monthly CPS public use files to generate monthly estimates of unemployment for homeowners and renters. We also use the monthly files when constructing our measure of mobility based on non-matches in the CPS. The source of the data is NBER and the US Census Bureau.[6]

3. Panel Study of Income Dynamics (PSID) - We use the PSID to calculate ($\frac{m_t^{UH}}{m_t^H}$) and ($\frac{m_t^{UR}}{m_t^R}$), as explained below.[7]

We go through our key measures in details below.

B.1 Key Mobility Measures

We measure total as well as inter-county mobility rates by using the March CPS. Since our exercise is focused on 2006 and onwards, our inter-county mobility rates are not exposed to the criticism made by Kaplan and Schulhofer-Wohl (2012*a*) regarding changes in the imputation procedures of mobility. We therefore include imputed values in our analysis. The advantage of that is that it allows us to calculate true mobility rates by dividing the number of moves between t and $t+1$ by the relevant population at time t, and not at $t+1$ (this could be a problem if dropping imputed values, and the total weight of imputed observations is changing over time).[8] One exception is the data presented in Section 2 (and the accompanying Table A.1). As this section reviews long term trends, we do exclude imputed values from our analysis, and divide t to $t+1$ moves by $t+1$ population.

For our counterfactual exercise, we want to correct our measured mobility rates for the possibility that unemployed homeowners move more than all homeowners. We used the Panel Study of Income Dynamics (PSID) to construct the ratio $\frac{m_t^{UH}}{m_t^H}$, where m_t^H is mobility rate for all homeowners. Table A.2 provides the annual values of the ratio $\frac{m_t^{UH}}{m_t^H}$ for homeowners and renters in columns 1 and 2 respectively. Note that the PSID becomes bi-annual after 1997. As the table shows, the ratio fluctuates a lot from year to year. Keeping this in mind, we do not take the value of this ratio from

[4] See http://cps.ipums.org/cps/ and King et al. (2010) for more details.

[5] See the NBER page with CPS March data: http://www.nber.org/data/current-population-survey-data.html. This is only used for footnote 10 of this web appendix.

[6] See http://www.nber.org/data/cps_basic.html and http://dataferrett.census.gov/ respectively for more details.

[7] http://psidonline.isr.umich.edu/

[8] The projection used in Table 1, column 2, Panel B uses inter-county data from 2001 for the projection, therefore potentially sensitive to the imputation critique. The results for this column are similar if we do drop the imputed records.

1 year but rather average over the 5 highest years. Despite these caveat, it is interesting to note that while the highest value for the ratio in the last twenty years is 1.97, the ratio was consistently around 2 during the twin recessions of the early 1980's and during the subsequent recovery. Also, the ratio of mobility for unemployed renters to total mobility of renters is generally a lot smaller, with a highest value of 1.25 in 1999.

We use self reported reason for move to distinguish job related moves from other moves. We try to take the broadest measure for mobility for job-related reasons, classifying the following answers to the why moved question in the CPS as moving for a job: "new job or job transfer", "to look for work or job lost", "for easier commute", "other job-related reason". The measured shares of movers for jobs among movers is presented in Table A.3. The first column shows the ratio for homeowners, and the second column for renters.

Finally, Table A.4 provides the distribution of movers over reason for move for the year 2006 from the CPS. Columns 1 and 2 give the distribution for homeowners and renters for all movers. Columns 3 and 4 give the distribution for homeowners and renters for inter-county movers only. The first four rows represent the categories that we consider to be job-related moves. Job-related moves were only about 16% of all moves and about 30% of inter-county moves for homeowners.

B.2 Renters Reweighting

We construct a measure of renter mobility which accounts for the fact the renters are different than homeowners on observable characteristics. To do that, we first run probit regressions by year (weighted by CPS weights) of mobility indicator on a set of demographic characteristic, focusing only on the renters sample. We then use our estimates to predict the probability of move for homeowners, and aggregate up using the homeowners weights.

In the choice of demographic characteristic to include in our probit regressions, we follow loosely the set of demographic characteristics surveyed by Kaplan and Schulhofer-Wohl (2012*b*) in the context of the long term trend in mobility, and include age categories, education categories (and and interaction of age and education), marital status, number of labor force participants in the households, income percentiles and number of family members.

B.3 Mobility Based on CPS Non-match Rate

In section 7, we use a measure of mobility based on the non-match rate in the CPS. The CPS is an address based panel. The non-match rate of the CPS from month to month is therefore informative about mobility.

Each sample housing unit in the CPS is interviewed 8 times. When a housing unit is first picked in the CPS sample, it is interviewed for 4 consecutive months. This is followed by an 8 month break and then another 4 months of interviews. Therefore at any one point, there are housing units which might have had between 1 to 8 interviews. If the number of interviews that a housing unit has given so far is between 1 to 3, or between 5 to 7, then the housing unit should appear again in the next month of the sample. If it does not (or if it does but the age, sex, or race of the household members has changed, see for example Madrian and Lefgren (1999)), then we can classify the household or the individual staying in the housing unit as someone who has moved. This non-match based mobility measure has the advantage that it is monthly, and that homeownership and unemployment status are observed pre-move. The disadvantage is that other than mobility, sources of non-matches include non-response, mortality and recording errors.

C Standard Errors Calculation

We provide details on how we compute the standard errors reported in Tables 1 and 2 of the paper. We compute the standard error of $\frac{U_{2012m03}-U^C_{2012m03}}{LF_{2012,m03}}$, using an iterative delta method.

First define

$$A_0 = [U_{H,2007m03}...U_{H,2012m03}, LF_{2012,m03}, m_{2008} * j_{2008}...m_{2012} * j_{2012}, m^C_{2008} * j^C_{2012}...m^C_{2012} * j^C_{2012}]'$$

to be the vector of variables which enter our counterfactual exercise. $U_{H,2007m03}...U_{H,2012m03}$ are the levels of unemployment for homeowners for each month from March 2007 to March 2012, $LF_{2012,m03}$ is the level of the labor force in March 2012, $m_{2008} * j_{2008} ... m_{2012} * j_{2012}$ are the shares of job-related moves among unemployed homeowners measured from the March CPS from 2008 to 2012, and $m^C_{2008} * j^C_{2012} ... m^C_{2012} * j^C_{2012}$ are the counterfactual shares of job-related moves among unemployed homeowners.[9] Given the covariance matrix of A_0, we can apply the delta method to find the covariance matrix of

$$A_1 = [U^C_{H,2007m04}, U_{H,2007m04} ... U_{H,2012m03}, LF_{2012,m03}, m_{2008} * j_{2008} ... m_{2012} * j_{2012}, m^C_{2008} * j^C_{2012} ... m^C_{2012} * j^C_{2012}]'$$

as A_1 is just a function of A_0. Note that $U_{H,2007m03}$ has been replaced by $U^C_{H,2007m04}$ in A_1. Similarly, the covariance matrix of A_1 can be used to compute the covariance matrix of

$$A_1 = [U^C_{H,2007m05}, U_{H,2007m05} ... U_{H,2012m03}, LF_{2012,m03}, m_{2008} * j_{2008} ... m_{2012} * j_{2012}, m^C_{2008} * j^C_{2012} ... m^C_{2012} * j^C_{2012}]'$$

This iterative process can be repeated till we find the covariance matrix of

$$A_{60} = [U^C_{H,2012m03}, U_{H,2012m03}, LF_{2012,m03}, m_{2008} * j_{2008} ... m_{2012} * j_{2012}, m^C_{2008} * j^C_{2012} ... m^C_{2012} * j^C_{2012}]'$$

and then as a final step apply the delta method again to estimate the variance of $\frac{U_{H,2012m03} - U^C_{H,2012m03}}{LF_{2012,m03}}$.

In order to implement this procedure we need to first estimate the covariance matrix of A_0. We do this in two steps. First we estimate the covariance matrix of $A_U = [U_{H,2007m03} ... U_{H,2012m03}, LF_{2012,m03}]'$. We do this by running a stacked regression of the form

$$u_{it} = \alpha_t + \varepsilon_{it} \tag{C.1}$$

where i indexes individuals, u_{it} is a dummy which takes value 1 if individual i is an unemployed homeowner at time t. Using CPS weights, u_{it} is regressed on a set of dummies, where each dummy takes value 1 for one month only. The coefficients on these dummies are the estimates for α_t, the share of unemployed homeowners at time t. To also compute the covariance of $LF_{2012,m03}$, we add an additional set of observations for all individuals in March 2012 on the left hand side which take value 1 for people in the labor force. We add an additional dummy on the right hand side which takes value one for only these additional observations.

The CPS sample design requires clustering at the primary sampling units level (PSU) which is the size of a county or an MSA. In addition, when using estimates from consecutive cross-sections, one needs to take into account the rotating panel structure of the CPS. Unfortunately, the PSU is not part of the public use micro data. We therefore compute the variance of α_t by applying two way clustering. We allow for correlation within households over time by clustering at the household level. We also cluster at the MSA level (we assign all individuals in a state whose MSA is not identified into a single group by state), which is a more conservative than the actual PSU stratification of the CPS. We then apply the CPS weights to recover the covariance matrix of unemployment levels.

In addition to the covariance matrix of A_U, we need the covariance matrix of $A_M = [m_{2008} * j_{2008} ... m_{2012} * j_{2012}, m^C_{2008} * j^C_{2012} ... m^C_{2012} * j^C_{2012}]'$. For specifications that use renters mobility in the counterfactual, we first compute the covariance matrix of $[m_{2008} * j_{2008} ... m_{2012} * j_{2012}]$ by assuming that it is diagonal and estimate the diagonal elements by using the replicate weights provided with the March CPS. $\left[m^C_{2008} * j^C_{2012} ... m^C_{2012} * j^C_{2012}\right]$ is a function of job-related mobility of unemployed homeowners in 2007 $\left(m^H_{2007} * j^H_{2007}\right)$, and job-related mobility rates for renters from 2007 to 2012 $\left[m^R_{2007} * j^R_{2007}, ..., m^R_{2007} * j^R_{2012}\right]$. We estimate the covariance matrix of $\left[m^H_{2007} * j^H_{2007}, m^R_{2007} * j^R_{2007}, ..., m^R_{2007} * j^R_{2012}\right]$ by assuming that it is diagonal and computing the diagonal elements by using replicate weights and then apply the delta method to get the covariance matrix of $\left[m^C_{2008} * j^C_{2012} ... m^C_{2012} * j^C_{2012}\right]$. All the specifications which do not use renters as counterfactual, use only a single number m^C_t for all years, and also use $j^C_t = j_t$. We account for that in two ways: First, we only assume that the covariance matrix for $[m_{2008} ... m_{2012}, m^C_t]'$ is diagonal and calculate it with replicate weights. We then assign $cov\left(m^C_t, m^C_s\right) = var\left(m^C_t\right)$ for all years when constructing the covariance matrix for $[m_{2008} ... m_{2012}, m^C_{2008} ... m^C_{2012}]'$. Second, we use replicate weights to calculate a diagonal covari-

[9] Note that we include only the number of unemployed homeowners in the vector A_0 and not the total number of unemployed in each month. This is because $\frac{U_{H,2012m03} - U^C_{H,2012m03}}{LF_{2012,m03}} = \frac{U_{2012m03} - U^C_{2012m03}}{LF_{2012,m03}}$ in our exercise. Therefore, finding the variance of $\frac{U_{H,2012m03} - U^C_{H,2012m03}}{LF_{2012,m03}}$ is equivalent to finding the variance of $\frac{U_{2012m03} - U^C_{2012m03}}{LF_{2012,m03}}$ and the former is a function of number of unemployed homeowners only and not of the total number of unemployed in each month.

ance matrix for $[j_{2008} ... j_{2012}]'$. We then apply delta method to calculate the (non-diagonal) covariance matrix for $[m_{2008} * j_{2008} ... m_{2012} * j_{2012}, m^C_{2008} * j_{2012} ... m^C_{2012} * j_{2012}]'$.[10]

References

Kaplan, Greg, and Sam Schulhofer-Wohl. 2012*a*. "Interstate Migration Has Fallen Less Than You Think: Consequences of Hot Deck Imputation in the Current Population Survey." *Demography*, 49(3): 1061–1074.

Kaplan, Greg, and Sam Schulhofer-Wohl. 2012*b*. "Understanding the long-run decline in interstate migration." Federal Reserve Bank of Minneapolis Working Papers 697.

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.

Madrian, Brigitte C., and Lars John Lefgren. 1999. "A Note on Longitudinally Matching Current Population Survey (CPS) Respondents." National Bureau of Economic Research, Inc NBER Technical Working Papers 0247.

Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review*, 95(1): 25–49.

[10]We always impose the assumption of diagonal covariance matrix for the values which are calculated using replicate weights. This is not necessarily true in the data, however in many specification, there is no other way to calculate the full covariance matrix. We check to see what the impact of non-zero covariances is on our standard error estimates for the baseline specification. For our baseline specification (Table 1, panel A, column 1), we computed the entire covariance matrix for A_0, by including dummies for mobility (m_t), counterfactual mobility (m^C_t) and job related mobility (j_t) in the regression in equation (C.1) again allowing for two way clustering. We then compared the standard errors from this covariance matrix to that from the same matrix but with the off-diagonal mobility terms (including the covariance of mobility rates with unemployment rates) set to 0, and find that the effect of assigning the off-diagonal elements to zero on the standard errors is negligible. In particular, assigning the off-diagonal elements to zero as we do in the paper makes the standard error about 0.8% larger.

Table A.1: Homeowner and Renter Mobility over Time

	(1)	(2)	(3)	(4)	(5)	(6)
Mobility Measure:	All Moves		Inter-county Moves		All Moves - H.O. Correction	
	Level	Logs	Level	Logs	Level	Logs
Homeowners trend	-0.126***	-0.014***	-0.076***	-0.022***	-0.168***	-0.018***
	(0.022)	(0.002)	(0.007)	(0.002)	(0.018)	(0.002)
Renters trend	-0.382***	-0.011***	-0.199***	-0.016***	-0.217***	-0.006***
	(0.024)	(0.001)	(0.019)	(0.002)	(0.072)	(0.002)
Homeowners X current recession	-2.016***	-0.328***	-0.559***	-0.315***	-0.399	-0.095*
	(0.304)	(0.040)	(0.103)	(0.036)	(0.406)	(0.052)
Renters X current recession	-1.048**	-0.046***	-1.353***	-0.164***	-4.977***	-0.167***
	(0.506)	(0.017)	(0.369)	(0.038)	(1.301)	(0.042)
Tests (P-values)						
Equality of trends:	0.000	0.158	0.000	0.015	0.518	0.000
Equality of interaction with current recession:	0.107	0.000	0.043	0.006	0.001	0.286
Observation	60	60	60	60	60	60

Notes: The table reports the results from the regression of annual mobility rates on dummies for homeowners and renters (not reported) and the interactions of a time trend and a current recession dummy with homeowner and renter dummies, with no constant. Current recession dummy equals 1 starting 2007. Mobility rates are calculated as the share of homeowners and renters in the labor force at t+1 who move between t and t+1 based on the March CPS for three mobility measures: All Moves (Columns 1, 2); Only inter-county moves (Columns 3, 4); All Moves, correcting for post-move recorded homeownership (Columns 5, 6). The time period is 1980 to 2011, where 1984 and 1994 are excluded, since mobility is not reported in CPS for these years. Observations with imputed migration data were removed for columns 1-4. Robust standard errors are reported. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.2: MOBILITY OF UNEMPLOYED BY MOBILITY FOR ALL - PSID

	(1) Mobility of Unemployed by Mobility of All - Homeowners (%)	(2) Mobility of Unemployed by Mobility of All - Renters (%)
1981	2.20	1.09
1982	2.09	1.07
1983	2.06	1.01
1984	1.70	1.09
1985	1.66	1.05
1986	2.00	0.86
1987	0.89	1.03
1988	2.59	1.01
1989	1.49	1.20
1990	1.15	1.04
1991	1.09	1.17
1992	1.43	1.18
1993	1.97	1.17
1994	1.30	0.97
1995	0.82	0.99
1996	0.87	0.80
1997	0.61	0.80
1999	1.69	1.25
2001	1.17	1.16
2003	1.07	0.86
2005	1.00	0.76
2007	1.18	0.98
2009	0.86	1.10

Notes: From the PSID. Column (1) shows the ratio of mobility for unemployed homeowners to mobility of all homeowners. Column (2) shows the same for renters. Timing for each row is from t-1 to t

TABLE A.3: MOBILITY FOR JOB AS SHARE OF MOVERS

	(1) Move for job - Homeowners (%)	(2) Moving for Job - Renters (%)
2006	15.40	22.47
2007	17.65	24.76
2008	16.74	25.45
2009	14.73	21.76
2010	14.47	20.33
2011	16.35	22.42
2012	16.30	23.99

Notes: From the March CPS supplement using CPS weights. Columns (1) and (2) show the share of all movers who move for job related reasons among homeowners and renters respectively. Timing for all columns is from t-1 to t

TABLE A.4: DISTRIBUTION OF MOVERS OVER REASON FOR MOVE

	(1) Reason for Move, All Moves - Homeowners (%)	(2) Reason for Move, All Moves - Renters (%)	(3) Reason for Move, Inter-county - Homeowners (%)	(4) Reason for Move, Inter-county - Renters (%)
New job or job transfer	7.13	11.19	17.32	24.93
To look for work or lost job	2.03	1.93	4.42	4.21
For easier commute	3.60	4.99	5.08	6.19
Other job-related reason	2.64	4.36	3.44	6.99
Change in marital status	7.49	6.03	7.10	5.46
To establish own household	6.43	11.41	5.28	6.12
Other family reason	10.28	11.26	13.93	13.36
Retired	0.27	0.18	0.67	0.51
Wanted to own home, not rent	24.09	1.24	13.91	0.81
Wanted new or better housing	16.49	18.45	8.02	6.67
Wanted better neighborhood	4.56	3.92	4.95	2.52
For cheaper housing	3.59	7.39	3.78	3.83
Other housing reason	5.55	10.24	3.91	4.61
Attend/leave college	1.23	3.51	2.27	6.48
Change of climate	0.32	0.54	0.79	1.14
Health reasons	0.59	0.68	0.78	1.20
Other reasons	2.67	1.36	2.90	2.33
Natural disaster	1.04	1.31	1.43	2.64

Notes: From the March CPS supplement using CPS weights for moves between 2005 and 2006. Columns (1) and (2) shows the distribution of movers by the reason for moving for homeowners and renters respectively for all moves. Columns (3) and (4) shows the distribution of movers by the reason for moving for homeowners and renters respectively for inter-county moves only.The population is restricted to people in the labor force.