Federal Reserve Bank of Chicago

Call for Papers

2012 Conference on Bank Structure and Competition

Third Quarter and Fourth Quarter 2011

Economic perspectives

Economic perspectives

Economic Perspectives is published by the Economic Research Department of the Federal Reserve Bank of Chicago. The views expressed are the authors' and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

Economic Perspectives and other Bank publications are available at www.chicagofed.org.

chicagofed.org

ISSN 0164-0682

Announcement

Final Print Edition of *Economic Perspectives*

To our subscribers,

As part of our ongoing efforts to reduce costs and protect the environment, we have decided to discontinue printing issues of *Economic Perspectives*. We will continue to publish the journal online on a quarterly basis.

We encourage you to visit our website, www.chicagofed.org, to sign up for an online subscription to *Economic Perspectives*. Our email alert service will notify you whenever we publish a new article online.

Our website also features a complete archive of past issues of *Economic Perspectives*, as well as other Federal Reserve Bank of Chicago publications.

We thank you most sincerely for your interest in our work.

Regards,

Helen Koshy

Helen Koshy,
Managing Editor
Economic Research Publications
Federal Reserve Bank of Chicago

chicagofed.org

Contents

Third and Fourth Quarters 2011, Volume XXXV, Issues 3 and 4

Evaluating the role of labor market mismatch in rising unemployment

Gadi Barlevy

Introduction and summary

From the second half of 2009 through the end of 2010, the U.S. labor market witnessed a systematic increase in the rate of job openings while the unemployment rate remained essentially unchanged. Some have argued that, evidently, the problem in the labor market during this period was not that firms were reluctant to hire additional workers, but that, for whatever reason, firms seemed unable to find suitable workers to staff the positions they were trying to fill. By this logic, using monetary policy to encourage further hiring by firms would have been unlikely to drive down unemployment: If firms were already trying to hire but could not, why should policy actions that mainly serve to encourage even more hiring have any impact on unemployment? The unemployment rate did finally register a decline in late 2010 and early 2011—a development that may eventually render less acute the debate about the need for monetary policy to address the problem of high unemployment. Still, constructing a framework for interpreting such labor market patterns and their policy implications remains an important goal. This is especially true given that there have been other periods in which job vacancy rates seemed to rise without a commensurately large fall in unemployment, although those episodes were not as dramatic nor as long as the most recent one.

In this article, I show how the labor market matching function approach developed by Pissarides (1985) and Mortensen and Pissarides (1994) can be used to assess the validity of the proposition that recent trends in vacancies and unemployment necessarily point to a diminished role for monetary policy. More specifically, I show that this framework indeed suggests that an increase in unemployment without a commensurate decline in vacancies can be indicative of a labor market shock that monetary policy cannot offset. However, this framework can also be used to derive a bound on how much a shock of this type can affect unemployment. Applying these insights to the period of the Great Recession reveals that this type of shock by itself would lead to an unemployment rate of 7.1 percent, considerably lower than the unemployment rate during most of this period. The higher actual unemployment rate suggests that other types of shocks, which monetary policy may be able to address, must also be operating. Hence, the recent patterns in unemployment and vacancy data do not necessarily rule out an important role for monetary policy. Whether more expansionary monetary policy would have been beneficial is a question that is beyond the scope of this article. Nevertheless, the matching function approach frames this question in a potentially useful way—that is, as a question of why the value of taking on additional workers appears to be so much lower now than in normal economic times.

My article is organized as follows. I begin by describing the matching function approach, and then I show that the shocks that affect unemployment in this framework can be decomposed into two groups—those that affect the ability of firms to find and hire qualified workers and those that affect the value to a firm of taking on an additional worker. I next explain how the model can be used to predict how a shock to the ability of firms to hire, calibrated to match the facts on unemployment and vacancies during the Great Recession, affects unemployment. Using this result, I argue that the increase in unemployment due to this shock is much smaller than the actual increase in unemployment during this period, so a shock to the ability of firms to hire cannot by itself account for the

Gadi Barlevy is a senior economist and research advisor in the Economic Research Department at the Federal Reserve Bank of Chicago. *The author thanks Dan Aaronson, Dale Mortensen, Ayşegül Şahin, Dan Sullivan, and Marcelo Veracierto for their helpful comments.*

rise in unemployment during this time. I conclude with a discussion about how measurement issues are likely to affect these conclusions.

The matching function approach

In this section, I lay out the key features of the labor market matching framework developed in Pissarides (1985) and Mortensen and Pissarides (1994). This framework rests on two key assumptions.

The first key assumption is that the total number of new hires h in any given period can be expressed as a function of the number of workers who are unemployed during that period, u, and the number of vacancies firms post over that same period, v:

1) $h = m(u,v)$.

This assumption is similar to the assumption invoked by macroeconomists that one can use an aggregate production function to express total output produced in a given period as a function of the total number of hours and the aggregate capital stock for that period. That is, the process by which unemployed workers looking for jobs and employers with vacancies looking for workers form new hires is assumed to operate with such regularity that one can reliably predict the number of new hires per period by using only data on the number of unemployed workers and the number of vacancies firms post. Empirical analysis supports the idea that the number of new hires can be related to the number of unemployed and vacant positions in a fairly predictable way. Much of this evidence is summarized in the survey by Petrongolo and Pissarides (2001).[1]

The function m is referred to as a matching function. A common restriction on the matching function is that the number of new hires h falls short of both the number of unemployed u and the number of vacant positions v. That is, some unemployed workers and some positions will remain unmatched by the end of the period. This is meant to capture various frictions in the process of filling new jobs from the ranks of the unemployed—such as a lack of coordination that leads multiple workers to apply to the same vacancies while other vacancies remain unfilled, or the fact that workers and firms do not initially know how well suited they are for each other and figuring this out can be time-consuming. Studies that explore these frictions in detail reveal that they do not always give rise to empirically plausible matching functions, and in some cases they suggest different interpretations for why hiring, vacancies, and unemployment are related. Moreover, these frictions sometimes imply that the number of new hires should depend on other variables besides just the number of unemployed workers and the number of vacant positions.[2] However, Petrongolo and Pissarides (2001) argue that the matching function approach performs quite well empirically and is suitable for analyzing certain questions concerning the labor market, just as assuming an aggregate production function is often useful for analyzing macroeconomic questions. That is, many macroeconomists are willing to posit an aggregate production function that is invariant to various shocks that affect the economy, even though the conditions under which one can ignore the decisions of individual firms in different sectors and express aggregate output as a function of aggregate inputs are quite stringent.[3] In defense of this assumption, these macroeconomists would argue that the aggregate production function performs well empirically, so it is likely to be useful in predicting how the economy would respond to shocks that only affect the aggregate capital stock and labor hours—for example, a change in income taxes that affects how much labor is supplied but does not affect the technology available for producing goods and services. Likewise, advocates of the matching function approach view the close empirical relationship between aggregate hiring and aggregate unemployment and vacancies as justification for ignoring the decisions of workers and employers that underlie the process of job creation. These advocates proceed as if hiring can be summarized with a mapping of aggregate vacancies and unemployment to aggregate new hires that is "structural," meaning that the mapping is invariant to shocks that affect unemployment and vacancies but not the frictions inherent in the matching process.

Petrongolo and Pissarides (2001) argue that the matching function is particularly well approximated by a Cobb–Douglas specification, that is,

2) $m(u,v) = Au^{\alpha} v^{1-\alpha}$,

where A is a scale parameter that determines the productivity of the matching process and α reflects the sensitivity of the number of new hires to the number of unemployed workers. That is, this specification will produce reasonably good predictions for the actual hiring rate given the unemployment rate u and vacancy rate v for coefficients α and A that remain stable over relatively long periods and that can be estimated from historical data.

The second key assumption of the labor market matching framework is that firms post vacancies as long as doing so remains profitable, implying that the expected discounted profits to a firm from posting a vacancy should be zero in equilibrium. Let J denote the value of a filled job to the employer who creates it, and let k denote the cost of posting (and maintaining) the vacancy, including screening and interviewing potential candidates. Then the assumption that employers are free to enter the labor market and attempt to hire workers as long as it remains profitable to do so implies that the value of a filled job times the probability of filling it should equal the cost of posting the vacancy, ensuring expected profits are equal to zero. Pissarides (1985) and Mortensen and Pissarides (1994) assume that each posted vacancy is equally likely to be matched, so with $m(u,v)$ new hires, the probability of filling a job is equal to $m(u,v)/v$. In this case, the implications of free entry can be summarized as follows:

3) $$\frac{m(u,v)}{v}J = k.$$

Not all models that give rise to a matching function representation as in equation 1 imply that the probability that any given firm expects its vacancy to be filled within the relevant period corresponds to $m(u,v)/v$. For example, even when there are underlying differences in firms, such as differences in the costs of processing applicants, we might still observe a stable relationship between aggregate hiring, unemployment, and vacancies. But different firms will assign different probabilities to filling their positions within the relevant period. Still, as long as $m(u,v)/v$ reasonably captures the probability of filling a position for the marginal firm at any point in time, proceeding with this assumption will be appropriate.

Substituting the Cobb–Douglas specification for $m(u,v)$ from equation 2 into the free-entry condition as given by equation 3 reveals that the free-entry condition can be expressed solely in terms of the ratio v/u.[4] This ratio is known in the literature as market tightness, since it reflects how many vacant positions are competing for each unemployed worker. In particular, the free-entry condition as given by equation 3 can be written as

4) $$A\left(\frac{u}{v}\right)^{\alpha} J = k.$$

As long as the parameters α and A remain constant over time, the free-entry condition as given by equation 4 tells us that if the value of a filled job relative to the cost of posting a vacancy, J/k, varies over time for any reason (such as a change in aggregate demand or changes in aggregate productivity), the market tightness ratio, v/u, would have to change as well. The fact that v/u changes with J/k ensures that firms continue to expect zero cumulative discounted profits from posting additional vacancies. Hence, if we knew how a particular macroeconomic event affected the ratio J/k, we could use the free-entry condition as given by equation 4 to deduce how this event should change the ratio of the vacancy rate to the unemployment rate we observe in the labor market. The two assumptions—the existence of a matching function and free entry into the labor market—thus impose a lot of structure on how various shocks affect labor market tightness as reflected in the ratio v/u.[5]

The Beveridge curve

With the introduction of one additional assumption, the labor market matching framework can be used to predict not only how v/u changes with J/k, but also how u and v change individually. In particular, suppose that the rate at which employed workers separate from jobs into unemployment is constant over time. This assumption may seem implausible at first, especially given the incidence of mass layoffs during recessions. However, the job separation rate that I need to assume is constant does not involve one-off spikes of job destruction that reflect immediate adjustment by firms to changes in economic conditions. Rather, the relevant rate is the one that corresponds to what happens in a recession once all bursts of job destruction are done.[6] Shimer (2005a) and Hall (2005) argue that fluctuations in this separation rate contribute little to overall changes in unemployment and can be ignored. In subsequent work, others argue that the separation rate appears to be quite cyclically sensitive, and find the separation rate makes an important but still relatively small contribution to overall fluctuations in unemployment.[7] However, their papers all look at the role of flows of workers from employment and unemployment without accounting for spikes of job destruction. Flows into unemployment that include bursts of job destruction may account for fluctuations in total unemployment, even if the separation rate that is relevant for my analysis is fairly stable. Later, I argue that data on unemployment and vacancies from three distinct episodes of high unemployment support the claim that the relevant separation rate, s, does not rise much during recessions. Moreover, if the separation rate were in fact higher during recessions, my calculation would only exaggerate the role of labor market mismatch, and the bound I derive for the effect of a shock to the ability of firms to hire would be too high.

To see what the model predicts for the behavior of v and u as opposed to their ratio v/u, consider what happens if J/k varies over time. Conditional on a given value of J/k, the free-entry condition as given by equation 4 tells us that the vacancy-to-unemployment ratio, v/u, must remain constant as long as J/k is constant. However, u and v could themselves change even while

J/k remains fixed, as long as they change in the right proportion. Still, one can show that as long as J/k remains fixed, u and v will converge to some steady-state values that depend on J/k and, moreover, that this convergence will be rapid. This quick pace of convergence is not just a theoretical result; rather, it has been confirmed empirically.[8] This finding may seem odd at first, since time-series data suggest unemployment is fairly persistent over time. However, it is important to note that I am referring to *conditional* (as opposed to *unconditional*) convergence in u and v. In other words, given a value of J/k, both u and v converge quickly to the steady-state values associated with this particular value of J/k. But if J/k follows a persistent process, unemployment will still appear to change slowly over time. Rapid *conditional* convergence is thus fully consistent with unemployment appearing to be a slow-moving process. Given that convergence to a steady state for a given J/k is quick, it follows that whatever the value of J/k happens to be at any point in time, the values of u and v we would observe should roughly coincide with the steady-state levels of these variables for that J/k.

To compute the conditional steady-state unemployment for a given J/k, note that the flow into unemployment is equal to $s(1-u)$, where s denotes the separation rate into unemployment, while the flow out of unemployment is equal to the number of new hires, $Au^{\alpha} v^{1-\alpha}$.[9] Since flows into and out of unemployment are equal in steady state, I can use this equality to arrive at an implicit formula for the conditional steady-state unemployment rate associated with a particular v/u ratio, which is associated with a particular value of J/k:

$$5)\quad u=\frac{s}{s+A(v/u)^{1-\alpha}}.$$

Rearranging equation 5 allows me to express the vacancy rate v implied by the model for a given unemployment rate u as follows:

$$6)\quad v=\left[\frac{s}{A}\left(u^{-\alpha}-u^{1-\alpha}\right)\right]^{\frac{1}{1-\alpha}}.$$

As long as the separation rate into unemployment s is constant, equation 6 implies a negative relationship between u and v. This relationship, when displayed graphically as a plot of the vacancy rate against the unemployment rate, is known as a Beveridge curve after the British economist William Beveridge, who first documented the negative relationship between the two series.

The negatively sloped Beveridge curve can be seen by plotting out the u and v implied by the model for different values of J/k. In particular, according to equation 4, changes in J/k will force changes in the ratio v/u. Intuitively, as jobs become more valuable, the probability of filling a job must fall to ensure firms still expect to earn zero profits. One can then deduce, from equation 5, that higher values of the ratio v/u imply lower values of u and, from equation 6, that lower values of u imply higher values of v.

Indeed, the only thing that induces a movement along a Beveridge curve as defined by equation 5 is a change in J/k. This result holds because the Beveridge curve in equation 5 is defined as the relationship between u and v for fixed values of A, α, and s. When these values are fixed, it is apparent from equation 5 that the unemployment rate u only changes if the ratio v/u changes. But when A and α are fixed, the free-entry condition as given by equation 4 tells us that the ratio v/u is entirely determined by J/k. Thus, a movement along the Beveridge curve occurs if and only if the value of taking on an additional worker relative to the cost of posting a vacancy changes. Various events can shift this value, including a change in worker productivity, a change in the bargaining power of workers, a change in aggregate demand, and a change in the employer's operating cost (such as a change to the cost of borrowing). But, for our purposes, all of these events can be grouped into a catchall category of shocks that affect the net value of a filled job or, alternatively, shocks that move the economy along a stable Beveridge curve.

The natural counterpart to shocks that induce a movement along a Beveridge curve are shocks that shift the Beveridge curve itself. As evident from equation 5, which defines the Beveridge curve, as long as s is fixed, the only way for the Beveridge curve to shift is if the matching function $m(u,v)$ itself somehow changes. A shift in the Beveridge curve thus corresponds to a shock that changes the way in which workers and employers come together to form new hires. One example of such a shock is a disruption that gives rise to greater mismatch between the skills employers require to fill their positions and the skills that unemployed workers currently possess—such as a shift in demand away from products the labor force is already skilled at making. Such a shock would presumably result in fewer positions being filled given the same number of unemployed workers and vacant positions, and thus, the productivity term A in the matching function would decline. The model thus delivers a clean dichotomy: Shifts of the Beveridge curve correspond to shocks to the ability of firms to hire (that is, changes in A), while movements along a fixed Beveridge curve correspond to changes in the incentives for firms to hire (that is, changes in J/k).

We can now recast the debate on the role of monetary policy in the face of high unemployment, using the terminology of the matching function approach. The observation that vacancies rose while unemployment was virtually unchanged implies the Beveridge curve must have shifted, that is, the hiring process became less efficient. There is arguably little monetary policy can do to affect the process by which firms and unemployed workers match up to generate new hires. However, whether there is any role for monetary policy depends on whether a shock to match productivity, A, is the only shock responsible for high unemployment. If the increase in unemployment is also due to a change in the relative value of a filled job, J/k, there may be some scope for monetary policy after all. So, for example, if the lower J/k reflects weak aggregate demand due to some underlying frictions, then monetary policy would have a role in addressing this. In the remainder of this article, I infer the decline in A from the shift in the Beveridge curve, and then I use the matching function approach to gauge how much a shock to A of this magnitude should have raised unemployment if the parameters that govern the value of a filled job remain equal to their levels during normal economic times (that is, to pre-recession levels). Since the implied unemployment rate falls far short of the actual unemployment rate that prevailed during this time, the high unemployment rate suggests that the relative value of a filled job, J/k, must have been lower during this period than during normal economic times. Whether this finding admits a role for monetary policy depends on why the value of a filled job is lower. Still, the calculation suggests that data on unemployment and vacancies do not rule out a role for policy per se and that high unemployment is due not only to an inability to hire among employers but also to a reduced willingness to hire.

Empirical Beveridge curves and estimating the matching function

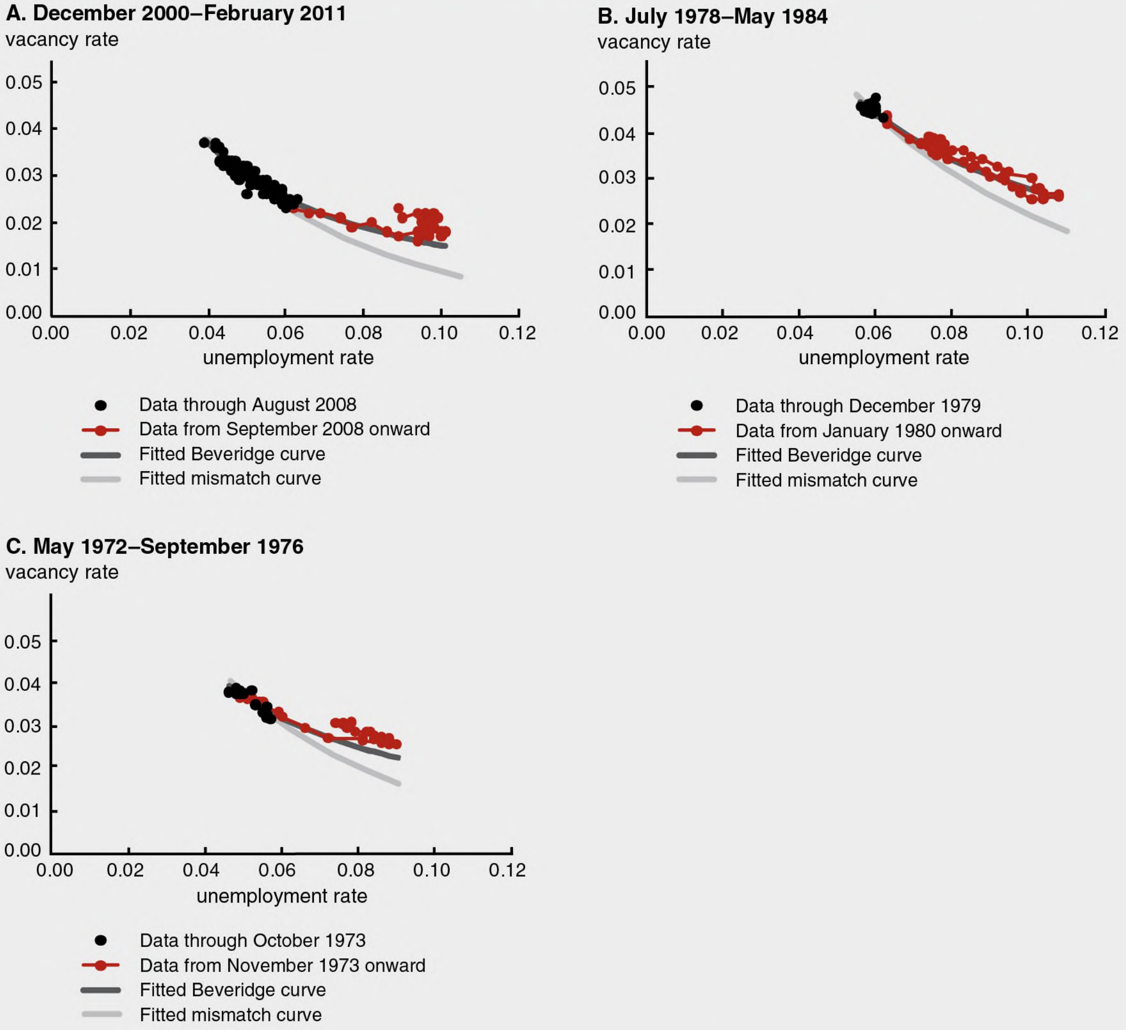
The first step in my analysis involves inferring the reduction in match productivity A over this period from shifts in the Beveridge curve. For this, I must begin with a benchmark value for A in normal times. I can do this by fitting the Beveridge curve relationship in equation 6 to data from before the recent crisis. That is, I estimate the parameters α and A of the matching function, using data only for the period before unemployment began to take off, and then I look at how this relationship holds up in predicting vacancy rates for observed unemployment rates once the unemployment rate begins to climb. To do this, I use data from the U.S. Bureau of Labor Statistics' *Job Openings and Labor Turnover Survey* (JOLTS), which begins in December 2000. To estimate the Beveridge curve, I use data through August 2008, just before the big run-up in unemployment that started a few months after the official start date of the recession according to the National Bureau of Economic Research (NBER). To estimate α and A, I follow Shimer (2005b)—who estimates the job separation rate at a monthly frequency—and set $s = 0.03$. However, the choice of s is essentially a normalization.[10] To infer A and α, I set out to match two specific aspects of the data. First, for each month, I use equation 6 to predict a vacancy rate v given an unemployment rate u in that month. The parameters A and α were chosen to ensure that the average predicted vacancy rate over all these months was equal to the actual average vacancy rate over the same period, namely, 0.029. Second, I chose the parameters to ensure that the difference in v between the start date and end date of my series, which is 0.013 in the data, matched the difference in the predicted v at these two dates. Matching the model and the data this way yielded values $A = 0.75$ and $\alpha = 0.46$. The implied (fitted) Beveridge curve corresponds to the dark gray line in panel A of figure 1, which is shown together with data on unemployment and vacancies. The points in black correspond to the data through August 2008 that were used to estimate the curve, while the points in red correspond to observations from September 2008 onward that were *not* used in estimating the curve. To help illustrate how u and v evolved from September 2008 onward, consecutive months are connected.

As evident in panel A of figure 1, the Beveridge curve implied by the model does a reasonable job initially of predicting the vacancy rate at each unemployment rate beyond the period it was estimated to match—in fact, beyond the historical range of both the unemployment and vacancy series used to estimate the curve. The forecast only starts to break down around August 2009, suggesting a change in the matching function. The fact that the curve fits well throughout the official recession as determined by the NBER—that is, from December 2007 through June 2009—and only breaks down afterward provides some reassurance that the separation rate, s, did not appear to rise significantly while the economy was contracting.

As a further check on how well the matching function approach fits the data, I went back and repeated the same exercise for two other periods with similarly high unemployment—namely, for November 1973–March 1975 and for January–July 1980 and July 1981–November 1982. Since JOLTS only begins in December 2000, I use the Conference Board's Help-Wanted Advertising Index for my measure of

FIGURE 1

Actual and predicted job vacancy rates for given unemployment rates during selected episodes



Notes: All curves are fitted only on data indicated by the black points. The fitted mismatch curves are based on Shimer (2007). See the text for further details.

Sources: Author's calculations based on data from the U.S. Bureau of Labor Statistics, *Job Openings and Labor Turnover Survey* and civilian unemployment rate series; and Conference Board, Help-Wanted Advertising Index, from Haver Analytics.

vacancies. This index is constructed using the number of newspaper advertisements for vacant positions. To transform this index into a vacancy rate, I normalized the series to coincide with the JOLTS vacancy rate for the period in which they overlap. For each recession, I followed a similar approach to estimating the matching function—that is, by taking data from a period prior to the recession to estimate the function and then seeing how the implied Beveridge curve does during the recession. However, since the Conference Board's Help-Wanted Advertising Index may be unreliable over long periods (given various gradual changes in the tendency of employers to rely on newspaper advertising for recruiting), I restrict attention to shorter periods for my estimation. For the 1973–75 recession, I look at the 18-month period before the NBER peak date, that is, May 1972 through October 1973. For the 1980 and 1981–82 recessions, I look at the 18-month period before the NBER peak date for the 1980 recession, that is, July 1978 through December 1979. In both cases, I estimate A and α in the same manner as for the data between December 2000 and August 2008—that is, I choose these parameters so that the average predicted vacancy rate over the period is equal to the actual average and the difference in vacancy rates between the start and end dates is the same in the predicted series as in the actual series. The estimated coefficients are reported in table 1, and the implied (fitted) Beveridge curves are illustrated as dark gray lines in the panels of figure 1 (the light gray lines are explained in the next paragraph). For the 1973–75 recession (panel C) and the 1980 and 1981–82 recessions (panel B), the data points used to estimate the curves are depicted in black, while the remaining data points are depicted in red. Note that in both panels B and C of figure 1, the original Beveridge curves are estimated from a period with little variation in the data, especially in the case of the 1980 and 1981–82 recessions. Still, the approach to using data before the recession(s) to estimate a curve performs well. For all of the recessionary periods I consider, the vacancy rate predicted for a given unemployment rate remains close to the actual vacancy rate once unemployment begins to rise. In both panels B and C of figure 1, the Beveridge curves do eventually appear to shift, although the shifts are much smaller than in the most recent episode, shown in panel A of figure 1. As evident in table 1, the coefficient α is estimated to be essentially the same in all three periods. This is consistent with my maintained approach of assuming the parameter α is fixed and that any changes in the matching function must therefore be attributed to A, the match productivity parameter.

For comparison, I also considered an alternative explanation for the matching function based on the notion of mismatch advanced by Shimer (2007), which was used by Kocherlakota (2010) to analyze the same labor market trends I consider here. The Shimer (2007) mismatch model offers a different interpretation for the relationship between new hires and unemployment and vacancies, and leads to a different zero-profit condition from equation 3.[11] Shimer's (2007) model also involves two parameters, which he denotes m and n. As I did earlier, I use the period before the recent run-up in unemployment (and the 18-month period before the start of the NBER recession for the two earlier episodes) to estimate these parameters and then consider how the model performs when unemployment rises. Following Kocherlakota (2010), in each case I choose these parameters to match the average values of u and v in the earlier period that is meant to reflect normal economic times. The estimates for the two parameters in each of the three episodes are summarized in table 2, and the implied curves relating unemployment and vacancies are shown in the respective panels in figure 1 in light gray. In all three episodes, Shimer's (2007) mismatch model predicts that the vacancy rate should decline more rapidly with unemployment than either what my estimated Beveridge curves predict or what we actually observe in the data. Since a shift in the curve in Shimer's model can be thought of as a shock to the ability of unemployed workers and job vacancies to match up with one another, this would suggest that all three recessionary periods and their subsequent recoveries were associated with significant rises in mismatch. While this reading of the data is certainly possible, it is striking that much of the discussion of the role of labor market mismatch during the Great Recession has tended to treat this phenomenon as exceptional; many of the explanations for the rise in mismatch in the labor market over the course of the Great Recession have emphasized features that are unique to this episode, such as the unprecedented collapse in house prices. Such views seem at odds with a specification that implies all three recessionary periods were associated with similarly large increases in labor market mismatch. The matching function approach is therefore more consistent with the view that the most recent episode is exceptional.

TABLE 1

Estimated parameters for a Cobb–Douglas matching function

	A	α
May 1972–October 1973	0.68	0.42
July 1978–December 1979	0.56	0.43
December 2000–August 2008	0.75	0.46

Note: See the text for further details.
Sources: Author's calculations based on data from the U.S. Bureau of Labor Statistics, *Job Openings and Labor Turnover Survey* and civilian unemployment rate series; and Conference Board, Help-Wanted Advertising Index, from Haver Analytics.

TABLE 2

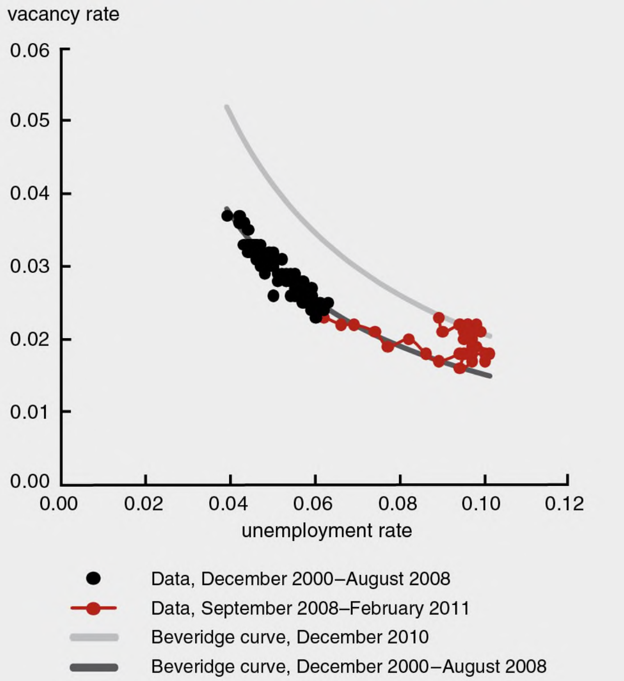
Estimated parameters for the Shimer (2007) mismatch model

	m	n
May 1972–October 1973	168.7	165.8
July 1978–December 1979	119.7	118.0
December 2000–August 2008	210.5	205.5

Note: See the text for further details.
Sources: Author's calculations based on data from the U.S. Bureau of Labor Statistics, *Job Openings and Labor Turnover Survey* and civilian unemployment rate series; and Conference Board, Help-Wanted Advertising Index, from Haver Analytics.

FIGURE 2

Fitting Beveridge curves during the Great Recession



Note: See the text for further details.
Sources: Author's calculations based on data from the U.S. Bureau of Labor Statistics, *Job Openings and Labor Turnover Survey* and civilian unemployment rate series, from Haver Analytics.

Inferring the extent of mismatch

After estimating the parameters associated with the Beveridge curve for normal economic times, I next turn to how the decline in match productivity A can be inferred from the apparent shift in the Beveridge curve following the most recent recession. Using data through August 2008, I know that the initial productivity of the matching function is given by $A_0 = 0.75$. I can deduce the value of A_1 needed to match a given unemployment and vacancy pair at any other point in time by using equation 6. For example, to match the data for December 2010, when $u = 0.094$ and $v = 0.022$, match productivity A_1 must solve

$$7)\quad 0.022 = \left[\frac{0.03}{A_1}\left(0.094^{-0.46} - 0.094^{0.54}\right)\right]^{\frac{1}{0.54}}.$$

Solving for A_1 yields $A_1 = 0.633$, that is, by December 2010 the productivity of the matching function declined 16 percent from its original level before the recession. Figure 2 shows the Beveridge curves for both values of A_0 (the dark gray line) and A_1 (the light gray line). In principle, I can fit a new Beveridge curve through any data point. The most recent observation at the time of this writing, for February 2011, lies on the same Beveridge curve implied by the data from December 2010, as evident in figure 2. Moreover, this curve is close to the highest curve one could fit through any of the data points between September 2008 and the end of the JOLTS sample. This leads me to focus on the curve that runs though the data point corresponding to December 2010 in measuring the decline in match productivity A.

An alternative way to infer the change in A over the course of the Great Recession would be to bring in additional data on new hires rather than only rely on the data for unemployment and vacancies. The idea is as follows. Since $m(u,v)$ corresponds to the number of new hires, which is measured in JOLTS, I can take the number of new hires and divide by the expression $u^{\alpha}v^{1-\alpha}$, using my previous estimate of $\alpha = 0.46$. In principle, this should give me a time series for match productivity A. This implied time series is depicted in figure 3.

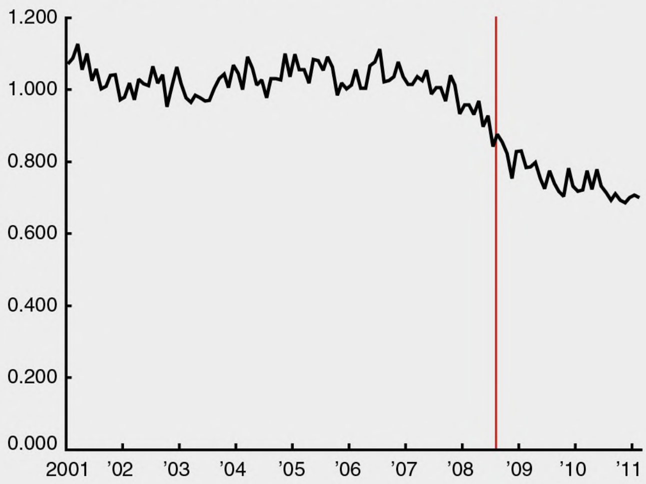
If I consider the period between August 2008 and December 2010, the implied match productivity declined by about 20 percent—a little larger than what I get without using hiring data and looking only at the implied shift in the Beveridge curve. However, as evident from figure 3, match productivity using data on new hires starts to fall around December 2007, considerably before any indications of a shift in the Beveridge curve relating unemployment and vacancies. The decline in match productivity between December 2007 and December 2010 is thus much larger, on the order of 25–30 percent. However, this decline is sensitive to the value of α, and it corresponds to 20 percent if I set $\alpha = 0.40$ instead of 0.46. But regardless of the precise value for α, data on hires suggest the decline in match productivity begins much earlier than the shift in the Beveridge curve. That said, both the magnitude and timing of the decline of matching efficiency depend on the measure of new hires used. Barnichon and Figura (2010) use the flows from unemployment to employment rather than all new hires, and find that the decline in A in 2009 is sharp, and its magnitude is comparable to what I estimate using the Beveridge curve.

Veracierto (2011) reviews several different approaches to estimating the productivity of the matching function, based on shifts of the Beveridge curve. These include accounting for flows into and out of nonemployment (see note 1, p. 94), measuring new hires based on flows into employment from either just unemployment or both unemployment and nonemployment, and using either shifts of the Beveridge curve or a comparison of changes in new hires to changes in unemployment and vacancies to deduce a time series for A. His preferred estimate suggests A had declined 15 percent since December 2007, in line with the estimate I infer from the shift in the Beveridge curve.

Since my calculations rely on the Beveridge curve specification in equation 5, I will use the estimate for the change in A based on how much the Beveridge curve shifted during the Great Recession in what follows.

FIGURE 3

Implied match productivity using data on new hires, 2001–11



Notes: The red vertical line corresponds to August 2008—the date at which the sample used to estimate the Beveridge curve in figure 1, panel A, ends. See the text for further details.
Sources: Author's calculations based on data from the U.S. Bureau of Labor Statistics, *Job Openings and Labor Turnover Survey* and civilian unemployment rate series, from Haver Analytics.

The effects of mismatch on unemployment

Once I determine that match productivity A declined by 16 percent between the level I estimate for normal economic times and the end of 2010, I can determine the effect of a shock of this size on the unemployment rate. To do this, I start at a steady-state unemployment rate of 5 percent, which roughly corresponds to the historical average of unemployment for the period covered by JOLTS through August 2008. From the Beveridge curve relationship implied by equation 6, I know the implied vacancy rate would have to be

$$v = \left[\frac{0.03}{0.75}\left(0.05^{-0.46} - 0.05^{0.54}\right)\right]^{\frac{1}{0.54}} = 0.03.$$

The implied ratio of u/v during these normal times will therefore equal $\frac{0.05}{0.03} = 1.67.$

Next, I use the free-entry condition as given by equation 4 to deduce how much a shock to A will affect the ratio u/v. To do this, suppose the shock to A had no effect on the ratio of the value of a filled job to the cost of posting a vacancy, J/k. In fact, J and k are determined endogenously, and changes in A can, and in many cases will, affect these values. However, for reasons I explain in more detail later, changes in A are likely to move J/k in a particular direction, implying that the unemployment rate holding J/k fixed will correspond to an upper bound on unemployment. Assuming J/k is constant thus offers a useful benchmark case.

Rearranging equation 4, I get $\frac{u}{v} = \left(\frac{k}{AJ}\right)^{1/\alpha}$.

Hence, given the estimated decline in the productivity of matching, holding J and k fixed, a decrease in A from 0.75 to 0.633 should lead the unemployment-to-vacancy ratio to rise by a factor of $\left(\frac{0.75}{0.633}\right)^{1/0.46} = 1.45.$

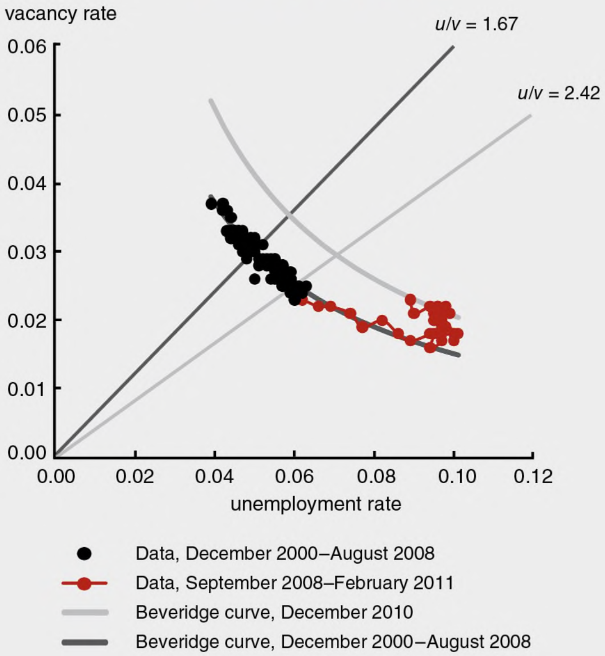
Given I needed the ratio u/v to equal 1.67 to support a 5 percent unemployment rate under the original Beveridge curve, I can deduce that the new equilibrium ratio of u/v will equal

$$1.45 \times 1.67 = 2.42.$$

Plugging in $u/v = 2.42$ into the Beveridge curve relationship in equation 5 when $A_1 = 0.633$ gives us the implied unemployment rate that must prevail in the new equilibrium:

FIGURE 4

Computing unemployment response from shock to match productivity *A*



Note: See the text for further details.
Sources: Author's calculations based on data from the U.S. Bureau of Labor Statistics, *Job Openings and Labor Turnover Survey* and civilian unemployment rate series, from Haver Analytics.

$$u = \frac{0.03}{0.03 + 0.633(2.42)^{-0.54}} = 0.071.$$

Thus, a shock to A, calibrated to the magnitude implied by the patterns observed in data on unemployment and vacancies alone, will raise the unemployment rate to 7.1 percent as long as it leaves the value of a filled job unchanged. Since 7.1 percent is much lower than the actual unemployment rate, this value suggests that shocks to the productivity of matching alone cannot account for the high unemployment rate.

Figure 4 illustrates the same calculation graphically. Each level of match productivity A is associated with a distinct Beveridge curve and a distinct ratio u/v determined by the free-entry condition as given by equation 4, which in the figure corresponds to the line emanating from the origin. The original Beveridge curve and free-entry condition associated with $A = A_0$ are shown in dark gray, while the new Beveridge curve and free-entry condition associated with $A = A_1$ are shown in light gray. A decline in A not only shifts the Beveridge curve but also rotates the free-entry condition clockwise to a degree that depends on the size of α. Intuitively, if hiring becomes less effective, firms will have an incentive to post fewer vacancies per unemployed worker, ultimately leaving more workers unemployed.

As I noted earlier, both k and J are in fact determined endogenously and will likely change when A does. For example, the process of creating a vacancy requires productive inputs such as labor, so the cost k will depend on wages that are determined endogenously. Since wages tend to rise and fall with economic activity both in the data and in the original Mortensen and Pissarides (1994) model, I would expect the cost of posting a vacancy k to fall as the unemployment rate rises. As for the value of a filled job to an employer J, there are various reasons to suspect it will be higher when there is more unemployment. Mortensen and Pissarides (1994) posit that the value of a filled job is determined as the result of Nash bargaining between workers and firms over the surplus from a match.[12] But the surplus from matching is higher when v/u is low, so a fall in A will lead to a higher value for J.[13] Intuitively, when it is easy to find a match, matching immediately is only slightly more valuable than separating and letting the two parties search for new matches, which they will likely find quickly. More generally, various realistic features that are absent in the benchmark model, such as curvature in the utility function and diminishing returns to labor, would tend to make a marginal job more valuable when fewer workers are employed. Essentially, diminishing marginal utility or diminishing marginal returns make another employed worker more valuable when fewer workers are employed. If both k falls and J rises at higher unemployment rates, the effect of a shock to A on u/v would only be smaller. As such, 7.1 percent should be viewed as an *upper bound* rather than a point estimate.[14] This result only reinforces the point that the high unemployment rate that was observed during this period should not be blamed solely on a decline in the ability of firms to fill their positions, but also on greater reluctance among firms to hire as reflected in a lower J/k.

Measurement issues

The calculations presented in the preceding section are based on the assumption that the inputs that go into creating new matches—namely, unemployment and vacancies—are measured accurately. However, there are reasons to suspect both series may systematically misrepresent the nature of inputs that enter the hiring process while the empirical Beveridge curve shifted. I now discuss each of these series in turn, as well as the implications of mismeasurement for my analysis. I will argue that in both cases, measurement issues only strengthen the conclusion that the decline in the ability of firms to hire cannot by itself account for the bulk of the increase in unemployment during this period.

I first consider the unemployment series. One distinguishing feature of the current episode of high unemployment is the exceptionally long duration of unemployment insurance (UI) benefits; in some U.S. states, the unemployed can receive UI benefits for up to 99 weeks. Indeed, several research papers have sought to estimate the effect of these extensions on both the unemployment rate and unemployment durations.[15] The extension of UI benefits can matter for my analysis in several ways, including the method by which unemployment is measured. First, though, it will be useful to review the various ways in which explicitly incorporating UI benefits into the model can matter for unemployment.

One reason UI benefits can matter is that they lower the cost of remaining unemployed, allowing workers to be more selective about which job they take. As a result, relative to the case in which UI benefits remained unchanged, unemployed workers will prefer to continue searching more often, and a smaller fraction of the contacts between unemployed workers and vacant positions will result in a match, that is, a new hire. Indeed, this provides one potential explanation for the apparent decline in match productivity. Note that this effect is already taken into account in the calculation I sketched out before; that calculation tells us how much a decline in the ability of firms to hire—for *whatever* reason—ought to affect unemployment. Indeed, all the papers estimating how much the extension of UI benefits contributed to unemployment find effects that are smaller than the bound I estimate.

Second, when a worker and an employer agree to form a match, the extension of UI benefits may require an employer to offer a worker higher wages given that more generous UI benefits improve the bargaining position of workers. This effect is emphasized in Kocherlakota (2011). Unlike the first effect that appeared as a lower value for A, this effect would show up directly as a lower value for J, the value of filling a job to an employer. Indeed, this may be one reason for why the value of a filled job to an employer appears to be lower now than it is in normal times.

Neither of these two effects poses a problem for determining whether the rise in unemployment can be attributed solely to a decline in the ability of firms to hire. Rather, they merely suggest potential interpretations for what might be driving shocks to A or J. However, there is a third potential implication of extending UI benefits that may act to distort measured unemployment and could pose a problem for my calculation. In particular, extended UI benefits may encourage disaffected workers who prefer to leave the labor force to present themselves as nominally unemployed in order to qualify for UI benefits. This will be the case even if such workers are not actively looking for a job beyond whatever token steps are needed to maintain their status. Such a phenomenon would make the measured unemployment rate seem higher than its true value. Formally, let u^* denote the fraction of the labor force that is actively looking for jobs, and let u_0 denote the fraction of the labor force that is not really looking for a job but reports itself as being unemployed. If the latter fraction literally takes no steps to search for a job, the matching process will only partner up the true unemployment and vacancy positions, and the number of hires will be given by

8) $h = m\,(u^*,v).$

At the same time, the official unemployment series will correspond to $u = u^* + u_0$, leading us to expect $m(u^* + u_0,v)$ hires. Since the matching function is increasing in both arguments, this will make the matching process appear less efficient than it truly is: We would observe surprisingly few hires given the seemingly large number of unemployed. Hence, the decline in match productivity A inferred from the shift in the Beveridge curve would exceed the true decline in A that enters the free-entry condition as given by equation 4. Since my approach provides an upper bound on the effect of a decline in the ability to hire on unemployment, though, overstating the decline in match productivity A will not overturn my results. If anything, it suggests the unemployment rate that should be expected from the decline in the ability of firms to hire is actually smaller than 7.1 percent.

Next, I turn to the time series for vacancies. Recent work by Davis, Faberman, and Haltiwanger (2010) has called into question whether vacancies provide a consistent measure of recruiting effort over time. In particular, they show that the vacancy yield, or the ratio of hires per vacancy, varies systematically across employers. For example, growing firms seem to be better at hiring, in the sense of being able to hire more workers per each vacancy posted. Davis, Faberman, and Haltiwanger (2010) argue that this pattern arises because the process of hiring requires firms to invest some effort into recruiting beyond posting the number of vacant positions they are seeking to fill.[16] They further reason that the same pattern should also occur over the business cycle: In recession times, when overall hiring is low, firms are likely to put in less effort into recruiting than in boom times. Thus, employers' hiring efforts would decline by more than would be reflected in the time series for the number of vacancies posted.[17]

Davis, Faberman, and Haltiwanger (2010) formalize these concerns as follows. Suppose that the effort that firms invest in recruiting can be summarized by the product of q and v, where q denotes recruiting intensity and v denotes the vacancy rate. The total number of hires is then given by

9) $h = m\,(u,qv)$.

That is, matching depends not on the number of vacancies, but vacancies together with how much firms invest in filling these vacancies. When recruiting intensity q falls below its historical average, the time series for vacancies v will fail to register this and will therefore overstate the overall recruiting effort. Using vacancies to proxy for recruiting efforts will then make matching efficiency appear to fall more than it in fact does. That is, we may wrongly conclude that firms find it more difficult to hire when in fact they are voluntarily choosing to search in a way that reduces the odds of hiring. Once again, this will cause us to overstate the decline in match productivity A from apparent shifts in the Beveridge curve and, therefore, to overstate the increase in the unemployment rate that can be attributed to less efficient matching now than in the past.

To provide a more quantitative illustration of this result, I can use the suggestion in Davis, Faberman, and Haltiwanger (2010) of proxying for recruiting intensity q by using the way in which the vacancy yield (hires per vacancy) varies across firms with different hiring rates. In particular, using variation in the vacancy yield across firms, they conclude that the elasticity of q with respect to overall firm hiring is given by 0.72. This implies setting $q = h^{0.72}$. Davis, Faberman, and Haltiwanger (2010) provide some evidence that this modification improves the time-series fit of the matching function. If this proxy is accurate, I can simply repeat the calculation for how much the apparent decline in match productivity should have increased unemployment, but replace the vacancy rate v in equations 1 through 6 with $h^{0.72}v$ as the second argument in the matching function.[18] Fitting a Beveridge curve to data on unemployment and this adjusted vacancy series through August 2008 yields $A_0 = 0.7$ and $\alpha = 0.54$. To match the data for December 2010 requires $A_1 = 0.605$, which is a smaller decline of only about 14 percent. For this decline, the implied unemployment rate due to just this shock to match productivity would be at most 6.3 percent. Correcting for measurement problems in vacancies can thus have a significant impact on how much unemployment is attributed to reduced effectiveness in hiring.

Conclusion

Recent labor market trends have raised concerns that the unemployment rate is high not because employers are reluctant to hire but because they are unable to hire—that is, for whatever reason, firms are unable to find suitable workers to staff the positions they are trying to fill. These concerns, if true, would cast doubt on using monetary policy to stimulate the labor market, since it works by encouraging firms to hire more. The matching function approach pioneered by Pissarides (1985) and Mortensen and Pissarides (1994) offers a framework for analyzing these issues. In particular, that framework can be used to separate the shocks that drive unemployment into two groups: shocks that affect the probability of finding a suitable worker and shocks to the value a worker generates once hired. The same framework allows us to estimate how much the probability of finding a worker declined and to compute a bound on how much this effect by itself would raise the unemployment rate. This bound as I have calculated it suggests that a decline in the ability to hire accounts for less than half of the total rise in unemployment during the Great Recession and that part of this rise in unemployment must be because firms find hiring less profitable.

While there is little monetary policy can do if firms find it more difficult to find suitable workers, there may be scope for monetary policy when firms find it less profitable to hire workers than during normal times. Whether such a role for monetary policy is warranted depends on why the value of a filled job to an employer is lower than in normal times. For example, if filled jobs are less valuable because of a shock that makes workers less productive, there is arguably little that monetary policy should do in response. But if jobs are less valuable because of insufficient aggregate demand on account of some market friction, there may be a role for monetary policy to stimulate demand. The key question for policy, then, is not what unemployment and vacancy data tell us about the possibility of mismatch, but why firms seem to find hiring workers less attractive than usual. Unfortunately, while the matching function approach is useful in pointing out the value of a filled job to an employer as an important variable, it offers little direct guidance as to why this value is so much lower now relative to normal times.

NOTES

[1]Of course, newly hired workers do not come only from the ranks of the unemployed; some were employed elsewhere, while others were not employed but did not report actively looking for a job either (that is, they were classified as "not in the labor force" by the U.S. Bureau of Labor Statistics, per the definition available at www.bls.gov/cps/cps_htgm.htm#nilf). In practice, the hiring rate can be accounted for quite well using data on unemployment, perhaps because the number of hires from out of the labor force and the number of hires of already employed workers move in opposite directions over the business cycle and tend to offset one another. One way to avoid the logical inconsistency of using data on unemployment to explain all new hires regardless of whether the worker was previously unemployed is to replace the number of new hires in equation 1 with the flow of workers from unemployment to employment, as in Barnichon and Figura (2010) and Veracierto (2011). While this approach restricts attention only to new hires who were previously unemployed, it suffers from the problem that the total number of vacancies is an imperfect measure of firm inputs into hiring the unemployed, since firms' efforts to fill these vacancies are aimed at hiring all workers and not just workers who are already unemployed.

[2]Petrongolo and Pissarides (2001) survey the microfoundations of the matching function, although several important papers in this area were published after their survey. The traditional model of coordination frictions, due to Butters (1977), assumes firms post vacancies, workers submit a single application each to some vacancy chosen at random, and each firm hires at random among the applications it receives. Burdett, Shi, and Wright (2001) emphasize that this model does not give rise to empirically plausible matching functions and that the number of hires per period will depend on additional variables, such as the size distribution of firms. Albrecht, Gautier, and Vroman (2003) assume workers can apply to multiple vacancies, but this does not give rise to empirically plausible specifications either. Lagos (2000) and Shimer (2007) model coordination frictions by letting firms and workers end up at different locations; firms choose locations at random and workers choose locations optimally to maximize their expected earnings (per Lagos) or at random (per Shimer). There are no frictions at any given location, so whichever side (firms or workers) arrives in smaller numbers winds up fully matched. Thus, each location will remain with either unemployed workers or vacant positions, but not both. Unemployed workers and vacancies are thus not inputs into forming new hires as the matching function approach implicitly assumes, but consequences of poor coordination between employers and workers on where to locate. When workers choose locations optimally, the matching function is not empirically plausible. When workers instead choose locations at random, the matching function matches the data well, at least for a certain range of unemployment and vacancies rates. Stevens (2007) develops a different theory of the matching function based on the notion that workers take time to screen heterogeneous jobs, rather than on coordination problems. She finds that the implied aggregate matching function is approximately Cobb–Douglas, as in equation 2 (p. 83). Decreuse (2010) develops a model where workers apply to jobs they do not realize are already filled. He finds that the implied matching function will depend on lagged variables beyond just the contemporaneous numbers of unemployed workers and vacant positions.

[3]For a survey that criticizes the use of aggregate production functions, see Felipe and Fisher (2003).

[4]The same is true more generally for any specification $m(u,v)$ that exhibits constant returns to scale.

[5]It should be noted that a recent body of literature, starting with Shimer (2005b), argues that the matching function approach suffers from serious shortcomings in its ability to match various labor market facts over the business cycle. However, this critique concerns whether the value of a filled job to the employer who creates it, J, varies enough over the cycle in these models, not whether the matching function can explain how new hires vary with unemployment and vacancies. My calculation does not depend on how J varies with aggregate conditions, nor does it impose much structure on how J ought to change over the cycle; and hence, it is not subject to this critique.

[6]More precisely, consider the Mortensen and Pissarides (1994) model where the separation rate into unemployment is endogenous. That model assumes jobs are hit with idiosyncratic shocks to the profitability of any given job at a constant rate λ per unit time. The shock term ε is drawn each time from some fixed distribution F. Firms optimally choose to terminate a job and send the worker into unemployment for severe enough shocks, that is, when ε falls below some critical level ε_d. Suppose that in a recession, firms become more demanding and raise the critical level to some higher value ε_d'. When the shock associated with the recession first hits, the unemployment rate will jump and the flow into unemployment will spike as all jobs whose ε lies between ε_d and ε_d' will be terminated immediately. The spike in the separation rate will appear large even when the regular flow into unemployment $\lambda F(\varepsilon_d)$ changes only modestly. My assumption that the separation rate is constant over time only requires that $\lambda F(\varepsilon_d)$ is relatively stable, not that flow rates from employment to unemployment (which will reflect spikes) be stable.

[7]Some examples are Mazumder (2007); Fujita and Ramey (2009); and Elsby, Hobijn, and Şahin (2010).

[8]In particular, Shimer (2005a) shows that the steady-state unemployment level to which the economy should be converging at any point in time can be readily computed from flows into and out of unemployment at that instant. He then shows that this steady state is nearly always close to actual unemployment.

[9]Barnichon and Figura (2010) and Veracierto (2011) also take into account flows between unemployment and not in the labor force in computing steady-state unemployment, which I ignore. Acknowledging that out of the labor force is a distinct labor market state does not change my ultimate conclusion that steady-state unemployment and vacancies will appear negatively related, although it may affect the shape of the curve relating the two series and how much we should conclude it may have shifted over time. I return to these issues later.

[10]Formally, as evident in equation 6, the Beveridge curve only depends on the ratio s/A. The levels of s and A depend on the frequency used to measure flows between labor market states.

[11]In particular, the probability of profitably hiring a worker in the Shimer (2007) model will not equal $m(u,v)/v$. Instead, it corresponds to the equilibrium fraction of locations with more workers than jobs. Employers in locations with more jobs than workers may still hire, but will earn zero profits. Although the probability of a profitable hire differs from $m(u,v)/v$, this probability will still be negatively related to v in equilibrium.

[12]Nash bargaining is one rule on how to divide a given amount of resources between two parties. This particular rule for how to divide resources was proposed by Nash (1950), who showed this rule had various desirable properties. Since employers and workers must divide the surplus that results from their joint production, Nash's solution has often been applied to determine the wage that workers receive.

[13]Kocherlakota (2011) shows that under Nash bargaining, J rises by nearly as much as A falls, so labor market tightness v/u is essentially the same regardless of A. In figure 4, keeping v/u unchanged but shifting the Beveridge curve up to the value associated with A_1 would

imply an unemployment rate of no more than 6 percent. But as hinted at in note 5, Nash bargaining is a somewhat problematic assumption, since for standard parameterizations it implies that productivity shocks produce fluctuations in J that are too small to explain business cycle volatility.

[14]In informal communication, Rob Shimer computed the effects of a 16 percent drop in the productivity of the matching function in a fully worked out equilibrium model with concave utility and declining marginal product of labor. He found that the unemployment rate would rise from 5 percent to 5.8 percent. This suggests my bound may be a substantial overestimate of the true effect.

[15]See, for example, Aaronson, Mazumder, and Schechter (2010); Valletta and Kuang (2010); Fujita (2011); Mazumder (2011); and Hu and Schechter (2011).

[16]More precisely, lower effort should be viewed as a change in some unobserved determinant of hiring that results in lower hiring rates for the same number of vacancies while holding unemployment fixed. This change may reflect lower effort—for example, firms may spend fewer resources on advertising a position or on screening and interviewing potential candidates. But alternatively, recruiters may raise the standards they expect from workers, which would also lower vacancy yields without representing lower effort on the part of the firms.

[17]Similarly, cyclical changes in the composition of which firms are trying to hire may lead us to incorrectly infer a change in the ability of the typical firm to hire. Suppose there was a rise in the share of hiring by firms that tend to rely more heavily on posting vacancies. In this case, measured vacancies would appear to rise more than overall hiring. Barnichon et al. (2010) provide evidence that the shift in the Beveridge curve coincided with a change in composition across industries toward industries that rely more heavily on posting vacancies to hire workers.

[18]Davis, Faberman, and Haltiwanger (2010) note that it may be difficult to ensure that both recruiting intensity varies over time and market tightness is determined by a free-entry condition such as equation 4. For example, they cite a model from Pissarides (2000, chapter 4) in which there is free entry into the labor market but recruiting intensity is constant over time. However, it is possible to get time-varying recruiting intensity in a model with free entry if we impose that both effective vacancies and the cost of creating effective vacancies to be homogeneous functions of q and v of the same degree. For example, since the product qv in equation 9 is homogeneous of degree 2 in q and v, the cost function for recruiting effort and posting vacancies must also be homogeneous of degree 2 in q and v.

REFERENCES

Aaronson, Daniel, Bhashkar Mazumder, and Shani Schechter, 2010, "What is behind the rise in long-term unemployment?," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 34, Second Quarter, pp. 28–51.

Albrecht, James, Pieter Gautier, and Susan Vroman, 2003, "Matching with multiple applications," *Economics Letters*, Vol. 78, No. 1, January, pp. 67–70.

Barnichon, Regis, Michael Elsby, Bart Hobijn, and Ayşegül Şahin, 2010, "Which industries are shifting the Beveridge curve?," Federal Reserve Bank of San Francisco, working paper, No. 2010-32, December.

Barnichon, Regis, and Andrew Figura, 2010, "What drives movements in the unemployment rate? A decomposition of the Beveridge curve," Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System, working paper, No. 2010-48.

Burdett, Kenneth, Shouyong Shi, and Randall Wright, 2001, "Pricing and matching with frictions," *Journal of Political Economy*, Vol. 109, No. 5, October, pp. 1060–1085.

Butters, Gerard R., 1977, "Equilibrium distribution of sales and advertising prices," *Review of Economic Studies*, Vol. 44, No. 3, October, pp. 465–491.

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger, 2010, "The establishment-level behavior of vacancies and hiring," National Bureau of Economic Research, working paper, No. 16265, August.

Decreuse, Bruno, 2010, "Matching with phantoms," GREQAM (Groupement de Recherche en Economie Quantitative d'Aix-Marseille), working paper, No. 2010-07, March.

Elsby, Michael, Bart Hobijn, and Ayşegül Şahin, 2010, "Disciplined estimates of gross flows among labor market states," presentation at Federal Reserve Bank of Chicago, September 10, available at www.econ.wisc.edu/workshop/Michigan.pdf.

Felipe, Jesus, and Franklin M. Fisher, 2003, "Aggregation in production functions: What applied economists should know," *Metroeconomica*, Vol. 54, No. 2–3, pp. 208–262.

Fujita, Shigeru, 2011, "Effects of extended unemployment insurance benefits: Evidence from the monthly CPS," Federal Reserve Bank of Philadelphia, working paper, No. 10-35/R, January.

Fujita, Shigeru, and Garey Ramey, 2009, "The cyclicality of separation and job finding rates," *International Economic Review*, Vol. 50, No. 2, May, pp. 415–430.

Hall, Robert, 2005, "Job loss, job finding, and unemployment in the U.S. economy over the past fifty years," *NBER Macroeconomics Annual*, Vol. 20, pp. 101–137.

Hu, Luojia, and Shani Schechter, 2011, "How much of the decline in unemployment is due to the exhaustion of unemployment benefits?," *Chicago Fed Letter*, Federal Reserve Bank of Chicago, No. 288, July.

Kocherlakota, Narayana, 2011, "Notes on 'Labor markets and monetary policy,'" supplementary material for speech at St. Cloud State University, March 3, available at www.minneapolisfed.org/news_events/pres/kocherlakota_notes_March3_2011.pdf.

__________, 2010, "Inside the FOMC," speech in Marquette, MI, August 17.

Lagos, Ricardo, 2000, "An alternative approach to search frictions," *Journal of Political Economy*, Vol. 108, No. 5, October, pp. 851–873.

Mazumder, Bhashkar, 2011, "How did unemployment insurance extensions affect the unemployment rate in 2008–10?," *Chicago Fed Letter*, Federal Reserve Bank of Chicago, No. 285, April.

__________, 2007, "New evidence on labor market dynamics over the business cycle," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 31, First Quarter, pp. 36–46.

Mortensen, Dale T., and Christopher A. Pissarides, 1994, "Job creation and job destruction in the theory of unemployment," *Review of Economic Studies*, Vol. 61, No. 3, July, pp. 397–415.

Nash, John F., Jr., 1950, "The bargaining problem," *Econometrica*, Vol. 18, No. 2, April, pp. 155–162.

Petrongolo, Barbara, and Christopher A. Pissarides, 2001, "Looking into the black box: A survey of the matching function," *Journal of Economic Literature*, Vol. 39, No. 2, June, pp. 390–431.

Pissarides, Christopher A., 2000, *Equilibrium Unemployment Theory*, 2nd ed., Cambridge, MA: MIT Press.

__________, 1985, "Short-run equilibrium dynamics of unemployment vacancies, and real wages," *American Economic Review*, Vol. 75, No. 4, September, pp. 676–690.

Shimer, Robert, 2007, "Mismatch," 2007, *American Economic Review*, Vol. 97, No. 4, September, pp. 1074–1101.

__________, 2005a, "The cyclicality of hires, separations, and job-to-job transitions," *Review*, Federal Reserve Bank of St. Louis, Vol. 87, No. 4, July/August, pp. 493–507.

__________, 2005b, "The cyclical behavior of equilibrium unemployment and vacancies," *American Economic Review*, Vol. 95, No. 1, March, pp. 25–49.

Stevens, Margaret, 2007, "New microfoundations for the aggregate matching function," *International Economic Review*, Vol. 48, No. 3, pp. 847–868.

Valletta, Rob, and Katherine Kuang, 2010, "Extended unemployment and UI benefits," *FRBSF Economic Letter*, Federal Reserve Bank of San Francisco, No. 2010-12, April 19.

Veracierto, Marcelo, 2011, "Worker flows and matching efficiency," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 35, Fourth Quarter, pp. 147–169.

Emergence of immediate funds transfer as a general-purpose means of payment

Bruce J. Summers and Kirstin E. Wells

Introduction and summary

In a modern economy, we pay for goods and services and trade in financial markets by transferring money held in accounts with banks. For the better part of the last century in the United States, most noncash payments were made with the paper check, a payment instrument that met most needs for payment services. Since the mid-1990s, use of the paper check has been in decline (Gerdes, 2008), a development that reflects technological advances and innovations by providers of payment services in response to needs for new and different payment instruments. Today, individuals, businesses, and governments can choose from a variety of payment instruments, each of which is designed to meet their specific needs for attributes such as certainty, speed, security, convenience, and cost (Foster et al., 2010). The most advanced means of transferring money between bank accounts is immediate funds transfer (IFT), which allows senders to pay receivers electronically in a highly convenient, certain, and secure manner, at low cost with no or minimal delay in the receivers' receipt and use of funds.

Today in the United States, IFT payments made through the banking system are mostly limited to large business transactions, interbank transfers, and specialized financial market transactions involving purchases of securities and the like. In total, these larger payments account for a small proportion of the total number of payments made throughout the economy. There is increasing evidence that the popularity of IFT is growing for everyday use, such as consumer purchases, payments between individuals, and small business accounts payable (Hough et al., 2010). To date, however, most general-purpose IFT payments are made on systems operated by nonbanks, the most familiar being PayPal.[1] The coverage of IFT systems supported by nonbank companies is limited to their closed customer groups, and transfers are made not in bank money but rather in special units of account defined by the nonbanks.

A notable development in a number of countries around the world is the everyday use of IFT for general-purpose payments using money held in accounts at banks. In these countries, banks have invested in applied technologies that allow them to provide low-cost IFT services to the general public, taking advantage of established national clearing and settlement arrangements that link all bank accounts together. As IFT innovators, banks in other countries are working together collectively and in cooperation with public authorities, such as central banks, to provide national clearing and settlement for the new IFT service.

This article examines the emergence of IFT as a general-purpose means of payment in the U.S. and in four other countries. We identify the public policy and business issues that arise when a new means of payment is introduced. We describe the attributes of payment instruments that users find attractive and compare the attribute profiles of different kinds of instruments, including IFT. We examine demand for IFT in the U.S. and present four international case studies of IFT. Finally, we discuss barriers to adoption of IFT in the U.S.

Payment attributes

Payments are made to satisfy personal or commercial obligations between and among individuals,

Bruce J. Summers is a consultant to the Financial Markets Group and Kirstin E. Wells is a business economist in the Financial Markets Group at the Federal Reserve Bank of Chicago. *The authors acknowledge and are grateful to the following for their helpful ideas and comments on earlier versions of this paper: Nancy Atkinson, Sujit Chakravorti, Harry Leinonen, Jeff Marquardt, Bruce Parker, Louise Roseman, the participants in a March 16, 2010, seminar at the Federal Reserve Bank of Chicago, and an anonymous reviewer.*

businesses (including nonprofits), and governments. Cash is the most basic and widely used means of payment by individuals in industrialized countries for transactions up to about $25 (Rysman, 2010; Smith, 2010). Apart from small-value payments, however, cash is not a preferred means of payment.[2] Most money is held in transaction accounts at depository institutions.[3] Payment instruments that provide access to this "deposit money," such as checks and debit cards, are the primary means of making payments (See box 1 for discussion of the bank payment business). Payment instruments are generally either credit transfers, whereby a payer (sender) directly authorizes the movement of money, or debit transfers, whereby a sender indirectly authorizes the movement of money via the payee (receiver). Regardless of payment type, the end result is the same; deposit money is transferred from sender to receiver.[4]

In the U.S., various payment instruments, supported by core processing systems in banks and interbank clearing and settlement mechanisms, are used to transfer deposit money. These include paper checks, payment cards, electronic debits and credits, and wire transfers (which, as we discuss later, are a specialized form of IFT). Senders select a payment instrument based on how well its attributes match the purpose of the payment (for example, point-of-sale transaction or trade payment between businesses). Because payments are two-sided transactions, the needs of both the sender and receiver are relevant in selecting the payment method to be used.[5]

The primary attributes considered by senders and receivers when selecting a payment instrument are as follows:

- Certainty—assurance to the sender and receiver that funds are transferred as ordered;
- Speed—timeliness of funds transfer from sender to receiver;
- Security—assurance that payment is protected against fraud and completed as ordered;
- Control—the sender and receiver have good information about and are able to control the timing of payment;
- Universal acceptance—the payment instrument is broadly accepted;
- Versatility—useful for a variety of personal and business transactions, including the ability to transmit remittance information; and
- Low cost and transparent pricing—reasonable cost relative to value; fees are clear to sender and receiver.

BOX 1

Transaction accounts and the payment line of business at banks

While considered part of the "payment business," bank transaction accounts offered to individuals and business customers are estimated to account for only a fraction of banks' total payment business revenue. Revenue from transactions accounts is attributable to net interest income earned from balances on deposit (typically the largest component), transaction fees, penalty fees, and a variety of other fees. The payment business also includes issuing credit cards to consumers, which is the largest piece of payment business revenue. Other payment businesses include issuing commercial cards, card services for merchants, money transfer services, issuing prepaid cards, and other smaller business lines (McKinsey & Company, 2009).

The fact that the majority of payment revenue does not come from transaction accounts, which are typically considered "core" banking services, can be explained by banks' ability to generate higher marginal returns from credit-related services. The transaction-account payment business has until recently emphasized "free" account services provided at very low fees, perhaps even below cost, as an inducement for customers to build accounts and grow net interest income.

Providers of payment services attempt to deliver these attributes in combinations that best meet the needs of the customers they serve. Technology is a principal catalyst leading to improvements in such services as one or more attributes can be strengthened without degrading other attributes.

A comparison of attributes across different payment instruments, including IFT, is shown in table 1, along with some common examples. Here, we discuss the attributes by type of payment instrument as summarized in table 1.

Payment types—Debit transfer

Debit transfers support the movement of money between accounts held with banks. Paper check and direct electronic debit are the most common debit transfer instruments. Historically, the paper check has been the most widely used method for making debit transfers. Paper checks have many attractive attributes, including payer control over the timing of payment and near-universal acceptance by payees. Checks are also very versatile in that they can be used for most personal, commercial, and government payments. Businesses in particular are heavy users of checks due to established back-office processes that link paper-based invoicing

TABLE 1

Attributes and examples of payment instruments

Attribute	Check	Direct debit	Direct credit	Wire transfer	General-purpose IFT[a]	Debit card
Certainty	Provisional payment to receiver	Provisional payment to receiver	Payment guaranteed to receiver	Payment guaranteed to receiver with immediate finality	Payment guaranteed to receiver	Payment guaranteed to receiver
Speed	Minimum one day	Minimum one day	Minimum one day	Real-time	Within minutes	Real-time authorization and guarantee; funds transferred end-of-day at the earliest
Security	Checks may be stolen and/or forged	Bank account and routing information from check can be used to originate debit transfer	Fraud is limited because payer directly sends funds from account	Fraud is limited because payer directly sends funds from account	Fraud is limited because payer directly sends funds from account	Card numbers may be stolen; use of PIN with certain cards limits unauthorized transactions
Control of timing	Payer controls instruction but cedes control of funds movement to payee	Payer controls instruction but cedes control of funds movement to payee	Payer controls transaction	Payer controls transaction	Payer controls transaction	Payer controls transaction
Universal acceptance	Yes	Sender and receiver must agree to use	Sender and receiver must agree to use	Yes	Closed system with limited number of users[b]	Limited by merchant acceptance
Versatility	Most types of payment transactions	Bill payments, business-to-business trade payments (with remittance information)	Recurring payments, business-to-business payments (with remittance information)	Financial market transactions (with limited remittance information)	Most types of transactions but limited POS	Point-of-sale (POS) and online only
Low cost and transparent pricing	Not transparent to individuals; per-transaction fee to businesses	Not transparent to individuals; per-transaction fee to businesses	Not transparent to individuals; per-transaction fee to businesses	High cost for sender and receiver (transaction fee)	Not transparent to individuals; ad valorem fee to merchant (PayPal)	Not transparent to individuals; ad valorem fee to merchant
Clearing & settlement	National check system	ACH system	ACH system	Accounts held with Reserve Banks or CHIPS	PayPal system	Card networks
Example	Business accounts payable	Utility payment	Payroll deposit	Purchase and sale of bank reserves	Purchase of goods and services	Grocery payment

[a]Information in this column is based on the features of PayPal, which is the nonbank IFT service most commonly used today by individuals (Shevlin, Fishman, and Bezard, 2010).
[b]As we discuss in the text, IFT in other countries links all or most transaction accounts held at banks.

and accounts payable systems to check-based payment systems. In general, the need to link remittance information with a payment is a key factor in a business's choice of a payment instrument and, historically, the remittance process has been paper-based.

For individuals, the cost of a check payment is not necessarily transparent because most banks bundle check fees with other transaction account fees. Some banks offer "free checking," which does not reflect the true cost. Businesses and governments are typically charged explicit per-item transaction fees by their banks, which, in combination with back-office processing costs, make checks relatively more expensive than electronic substitutes (Wells, 1996). Despite higher costs, many business users find established payment processes effective and the cost of switching to an electronic workflow, including persuading counterparties to accept electronic payments, prohibitive (AFP, 2010).

Historically, the process of clearing checks, which involves moving the check from sender, to receiver, to receiving bank, to paying bank (possibly through intermediary banks or a central clearinghouse), was labor and capital intensive. Today, checks are converted to digital images for electronic processing once they enter the clearing process. This may happen at a merchant location, even as early as the merchant's point of sale.[6] Even though most checks are cleared electronically, funds movement is still a relatively slow process. Depending upon when checks are entered into the collection process by the receiving bank, provisional credit is available to a receiver either the same day or the next day, and deposit money is transferred from the sender's bank within one or two days.

Another type of debit transfer is the electronic equivalent of a check, called direct debit. Direct debits are marketed to individuals as "autopay" or "direct bill." This instrument allows individuals to make payments directly from their bank accounts by supplying their bank account and routing number to the payee. The true cost of direct debit is hidden because it is typically free, or bundled with account service fees. Direct debits are used primarily to pay bills and, more recently, for online purchases. Acceptance of direct debit is limited because not all payees offer this option to individual payers.

Businesses are heavy users of direct debits to make and receive trade payments, because fees are lower than for checks and because electronic payments support greater back-office operating efficiency. Direct debits are typically as versatile as checks because remittance information may be included electronically with payments. Yet, acceptance is limited because both the sender and receiver must agree to use electronic payments.

Direct debits are cleared and settled via the automated clearinghouse (ACH) network, to which payees gain access through their account-holding banks. Payment transactions are sent in batch form to a central operator for processing with settlement at pre-scheduled times during the day. Sending and receiving banks subsequently update the accounts of senders and receivers. The ACH was designed as a batch system because checks are processed in batch form, and this processing model persists to this day. Because of batch processing, ACH debit transfers are relatively slow—there is a one-day gap between the time a payment is initiated and the time deposit money is transferred. Thus, direct debits, though electronic, are not necessarily quicker for end-users than check payments.

As shown in table 1, checks and direct debits fall short in terms of certainty, control, and security. Because payees initiate the movement of funds from the accounts of payers, payers are uncertain about the timing of the movement of funds. The lack of certainty and control for payers has a direct bearing on payment fraud, because someone who has obtained bank account and routing information from a stolen check, for example, may be able to initiate an account debit without a payer's knowledge by fraudulent means. Fraudulent payments, once identified by the payer or the payer's bank, may be returned, but returned payments undermine certainty and security.

Credit transfer

Credit transfer is accomplished in a variety of ways, principally as electronic credit and IFT.[7] Electronic credit transfers are used by businesses and governments to make recurring payments to individuals for obligations, such as payroll and social security payments. They are also used for business trade payments. Recurring payments are received by individuals as "direct deposit." Direct deposit is used for nearly all government-to-individual payments, but not all businesses have adopted direct deposit. The cost of direct deposits is not transparent to individuals because they are typically not charged to receive them, whereas business users pay an explicit per-transaction fee.

Direct deposits and some other types of electronic credit transfers are processed on the ACH network. As in the case of debit transfers, ACH credit transfers are relatively slow, with a one- or two-day lag between the time the payment is initiated by the sender and the time deposit money is transferred to the receiver. As shown in table 1, electronic credits offer more certainty, control, and security for senders, who directly authorize the movement of money.

Immediate funds transfer is used today primarily for large-value business and financial market transactions, through bank wire transfer services. Wire transfers constitute a small portion of the overall number of payments and a large portion of the overall value of payments; their daily value exceeds a trillion dollars. Wire transfers are expensive, typically costing about $25 to $35 per transaction, and are thus not widely used by individuals. Wire transfers are not only immediate, they are final. That is, wire transfers are irrevocable and unconditional and offer the highest certainty of any payment type. Wire transfers are accepted by most banks.

Clearing and settlement of wire transfers takes place over one of two specialized systems: Fedwire, which is operated by the Federal Reserve Banks, or the Clearing House Interbank Payment System (CHIPS), which is operated by The Clearing House Payments Company L.L.C. In the case of Fedwire, banks transfer balances directly between accounts they hold with the Federal Reserve Banks. CHIPS is a closed network whose members exchange payments, which are settled by means of continuous multilateral netting. As indicated in table 1, wire transfers are quick, certain, and secure, and accordingly they are relied on in interbank and financial markets worldwide and are often made using real-time gross settlement (RTGS) systems (World Bank Group, 2008). Virtually all RTGS systems, including Fedwire, are operated by central banks, which for these purposes are functioning as universal bankers' banks. Wire transfers involve the transfer of deposit money that banks hold in accounts with central banks (sometimes referred to as "central bank money"). Public oversight authorities have made the use of RTGS a virtual requirement for systemically important payment systems (BIS, 2001).

Much of the innovation in U.S. payment instruments over the past decade has centered on general-purpose IFT. Nonbanks have been at the forefront of this innovation. The approach taken by nonbanks is twofold: 1) offer payment services directly to end-users that substitute for and compete with the services provided by banks; and 2) provide banks with the business processes and technical capabilities that allow them to offer IFT services to their account-holding customers.[8]

Under the first approach, nonbanks directly provide general-purpose IFT services to individuals and small- to medium-sized businesses. A nonbank payment provider must first establish a funding source for IFT payments that are initiated by its customers, as it cannot tap directly into the customers' bank accounts. The nonbank provider would typically do so by setting up an omnibus account with its bank, to which its customers make deposits. The customer funds pooled in the omnibus account are then reflected in ledger accounts set up by the nonbank on its computers that are denominated not in commercial bank money, but in parallel units of value identified with the nonbank provider (for example, PayPal dollars). Collectively, these ledger accounts constitute a closed, proprietary network that supports transfers of value units among the users of the nonbank providers' services. Payments to receivers outside the network are supported, but in this case a conversion back to bank money is required. The conversion back to bank money is accomplished by sending deposits in the omnibus account back through the bank payment network to the bank account of the receiver, which is not part of the nonbank network. The nonbank payment networks rely on modern, applied technologies to support immediate funds transfers, and in-network transfers occur virtually instantaneously. Out-of-network transfers that rely on the banking system may take several days to complete.[9]

Under the second approach, banks use a technology platform supplied by the nonbank company in combination with their own in-house authorization systems to provide IFT services to their account-holding customers. Banks following this approach brand the services as their own. Again, however, the resulting network is closed, and proprietary, connecting accounts at the limited number of banks that use a particular nonbank vendor's platform. So long as a payee and payer hold accounts at banks that use the same nonbank provider's technology, they can transfer funds directly to each other's accounts.[10] Out-of-network transfers are possible, but again the transfer may take several days to be completed.

Debit cards

Debit cards are a unique type of payment. While payments made by debit card are cleared and settled like debit transfers, they offer IFT-type attributes to both cardholders and merchants, as shown in table 1. In particular, debit card payments offer speed, certainty, and control to both parties. Specialized authorization systems instantaneously check, at the point of sale, whether payers are able to fund purchases from their bank accounts. Once a transaction is authorized, merchants have the certainty of knowing that payment will be received. Unlike IFT, however, funds are not transferred from the individual's to the merchant's account until the end of the day at the earliest. Yet, the pre-authorization makes the payment seem immediate to cardholder and merchant.[11]

Debit cards offer limited versatility, as they are used primarily at the merchant point-of-sale, with merchants who have agreed to join a debit card network. The cost of debit cards is not transparent to cardholders (typically transactions are free), and merchants pay ad valorem fees, which are a percentage of the transaction amount.[12] Debit cards are subject to unauthorized use if stolen, and the card networks have security measures in place to limit unauthorized transactions as well as rules on limited liability for merchants. (Credit cards are not taken up directly because, as described in box 2, their principal purpose is to provide credit services.)

IFT innovation—General-purpose payments

The foregoing discussion of payment instruments and their attributes shows that wire transfer and general-purpose IFT offer attractive combinations of attributes compared with other types of payment instruments, especially certainty, speed, control, and versatility. The average price of a wire transfer makes this payment instrument unattractive for general-purpose use, and a primary advantage of IFT is its low price. As we discuss in the next section, evidence of latent demand and revealed preferences for certain combinations of attributes support the view that there is an unmet need for broadly available IFT in the U.S.

> **BOX 2**
>
> **Credit cards**
>
> Credit cards are also commonly used by individuals at merchant locations, yet credit card transactions are not debit or credit transfers. Credit cards are a means of providing access to short-term consumer finance, whereby merchants receive funds from their banks at the end of the day but cardholders do not authorize the transfer of deposit money until they pay their monthly credit card bill to the bank that issues them the card. This bill is for the aggregate amount owed to cover multiple transactions and is not required to be paid in full. Thus, credit card transactions, while often considered payment transactions, do not fall under either the credit transfer or debit transfer model. The distinction between a credit card transaction and payment transaction holds true even though an estimated 40 percent of cardholders, so-called convenience users, do not rely on short-term credit and pay their balance in full each month (Herbst-Murphy, 2010). Convenience users typically use credit cards for other reasons, such as garnering reward points or simplifying their cash management by accumulating payments over a monthly grace period.

Demand for IFT

Latent demand

Research conducted by the Federal Reserve System on payment system user preferences provides evidence that users desire a service with the attributes of IFT. In a 2002 survey on the future of retail electronic payments (Board of Governors of the Federal Reserve System, 2002), respondents appealed for the development of a low-cost way for individuals and businesses to make online real-time funds transfers.[13] Survey participants also noted the need for a new, uniform "deposit directory" of account numbers and account status, or some other means of account verification, as well as a directory to route electronic payments more easily to recipients. Further, in a 2006 survey on barriers to innovation in payments (Board of Governors of the Federal Reserve System, 2006), payment industry respondents indicated that wire transfers would be an effective mechanism for making smaller value payments at an acceptably low price (presumably the price would need to be lower than the typical bank wire transfer fees) and with remittance information easily linked to corporate billing systems. These two surveys reveal a clear interest in IFT, subject to the availability of directory and routing information and responsiveness to specific user requirements, including low cost and improved support for remittance information.[14]

Revealed preferences

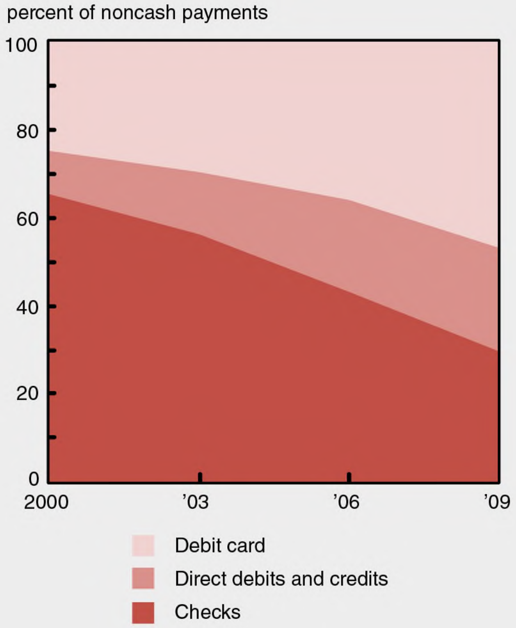
Other evidence to support the view that IFT may be broadly desirable in the U.S. is the increased use of payment instruments that offer attributes most closely related to IFT. For example, the use of debit cards, which offer more control, certainty, and speed than other payment instruments, has grown more rapidly than that of any other means of payment for point-of-sale and online purchases by individuals. In 2008, individuals held more debit cards than credit cards and, on average, used debit cards more often than cash, credit cards, or checks individually (Foster et al., 2010). In 2008, $1.00 of every $5.00 was spent with a debit card in the U.S., up from $1.00 of every $14.00 in 2001 (Herbst-Murphy, 2010).

A portion of the increase in debit card usage can be explained as a secular trend of growing familiarity with electronic payments in general. As shown in figure 1, the percentage of noncash payments made by electronic methods has grown in the last ten years, which reflects this trend. Other reasons cited for debit card preference include increased convenience and speed of payment (Rysman, 2009), which make debit cards more attractive than checks. Part of the growth in debit card usage and decline in check usage shown in figure 1 can be attributed to the substitution of debit cards for checks.

Business use of payment instruments with attributes that closely resemble those of IFT has grown as well. In 2010, one of the fastest-growing transactions processed on the ACH network was direct credit for

FIGURE 1

Rise in IFT-like payments



Source: Board of Governors of the Federal Reserve System (2010, 2007, and 2004).

sending bills paid through online banking sites to biller receivers (Digital Transactions, 2010). Direct credits offer advantages over checks and direct debits for bill payment in terms of certainty and security, much like an IFT.

Experience with IFT in other countries provides insights into the potential for this type of payment in the U.S. In the next section, we present four international case studies of the successful introduction of IFT. In each case, IFT has been introduced as a universal or near-universal payment instrument supported by clearing and settlement mechanisms that connect virtually all bank accounts within a given country. Universal support for IFT has been accomplished through industry-wide cooperation, sometimes facilitated and promoted by public authorities.

IFT case studies

As we noted earlier, wire transfer is a standard means of payment worldwide and is most often supported by RTGS systems operated by the central banks. These RTGS systems are capital intensive, benefit from economies of scale, and in most cases are operating well below efficient scale (Allsopp, Summers, and Veale, 2009). The services provided by RTGS systems in at least seven countries have been expanded to general-purpose payments. These countries are China, the Czech Republic, Serbia, the Slovak Republic, Switzerland, Turkey, and Ukraine.

The banking systems of at least three other countries have created transaction processing infrastructures specifically designed for IFT. These countries are Mexico, South Africa, and the United Kingdom (UK). Consequently, although their implementation approaches may differ somewhat, the banking systems of at least ten countries have taken coordinated steps to provide IFT services. Here, we discuss the cases of Mexico, South Africa, Switzerland, and the UK.[15] These case studies help us to identify several business and public policy considerations that arise when a country seeks to establish a national network to support a new payment instrument. A common consideration is reliance on the national RTGS system to provide finality for IFT payments, either directly by means of transaction processing or indirectly by means of interbank settlement of IFT obligations.

Mexico

Immediate funds transfer was introduced in Mexico in 2004, with the implementation of a new RTGS system by Banco de México. The new RTGS system, known by the acronym SPEI, takes advantage of new processing technologies that allow continuous upward scaling of transaction processing volumes at low marginal cost, with strong security based on a public key infrastructure (PKI). During the SPEI project, some commercial banks indicated that they considered two credit transfer systems (the other being the Mexican ACH) to be wasteful. Accordingly, Banco de México designed SPEI to support a variety of credit payments on one processing system, providing banks with a choice between using the new RTGS and ACH. Banco de México has promoted the use of IFT through advertisements in the mass media.

The central bank also provides payment services to the Mexican government and had been using its old RTGS for large government disbursements and the ACH for smaller disbursements. It was clear that so long as the Mexican government continued using the ACH for any disbursements, commercial banks would be forced to maintain their ACH systems. In 2008, the government agreed to Banco de México's request to use SPEI for all disbursements. Further, the government decided to centralize its payroll processing and use SPEI for government payrolls by the end of 2009. To support government payments, Banco de México instituted an earlier opening time for SPEI in order to allow commercial banks to maintain their established processing schedules. The government and banks use the straight-through processing capabilities that SPEI offers, with the expectation that both efficiency and service levels will increase throughout the payment system.[16] Most SPEI payments take less than a couple of minutes to reach the beneficiary's accounts. By law, all SPEI payments are final, regardless of their size or the beneficiary. Payments are final as soon as the beneficiary's bank receives a settlement notice.

Mexican commercial banks offer their customers IFT payment services mainly online. The payer must provide the bank routing and account numbers for the payee. One-off payments are therefore difficult to make because of the information that is needed on the payer side. Point-of-sale transactions are not currently supported, in part because of stringent security requirements established by the Mexican Banking Commission. Small mobile payments are, however, now being supported by new regulations and by a security agreement between banks and the commission.

Banks follow a variety of practices for pricing IFT payments. Large banks charge per-transaction fees of up to $0.35 or bundle credit transfer services with their Internet banking offerings for a fixed fee. The typical fixed fee for Internet banking service in Mexico is around $2.50. Prices for over-the-counter payments usually are higher than for Internet banking transactions. Some banks charge about half as much for ACH credit transfers as for real-time credit transfers, whereas other banks charge the same for both payment services.

South Africa

The introduction of IFT services for use by the general public in South Africa is a direct result of a recent initiative by commercial banks. The South African payment system has supported a number of general-purpose payment options, including the paper check, the check card (a means of initiating a credit transfer from a checking account at the point of sale, upon authorization, and usually available only to high-net-worth customers), debit and credit cards, and ACH-type electronic funds transfer (EFT) debit and credit payments. Access to check payments would take from one to seven days; and EFT and Internet payments would take on average one day for the transfer of funds intrabank and three days for the transfer of funds interbank.

Commercial banks in South Africa identified the need for a payment instrument that would give the general public the ability to transfer funds quickly and in a manner that made funds available to the payee immediately. Seven banks began collaborating in 2005 to develop a new clearing and settlement mechanism called Real-Time Clearing (RTC), in cooperation with the South African Reserve Bank, and the capability was implemented in March 2007. The banks provide services via Internet banking for consumers, online initiation through corporate banking solutions for businesses; and offline, over-the-counter initiation at a bank branch or by telephone. In each of these cases, the payer must follow an authentication procedure and provide routing information (bank and account number) for the payment. While no point-of-sale facilities are currently available, mobile services over cell phones are supported; and in theory, a merchant could be paid by mobile IFT, although no confirming message would be sent to the payee.

Immediate funds transfer payments made by the RTC method are governed by rules established by the Payment Clearing House (PCH), which banks are bound to in bilateral agreements. In addition to rule-making, the PCH functions as the system operator. It clears RTC payment instructions and provides the interface to the South African Reserve Bank RTGS system, known by the acronym SAMOS, which clears and settles the interbank obligations arising from RTC. Once an RTC payment instruction is cleared by the PCH, the receiving bank credits the beneficiary's account within 60 seconds. The interbank RTC clearing and settlement obligations built up in the PCH are sent to SAMOS on the hour every hour during the business day, which significantly reduces the risks associated with RTC payments.

Banks charge higher prices for IFT than for other Internet banking and mobile payments. Pricing has two parts, a per-transaction fee and a charge based on the amount transacted for purchases, with a cap on the maximum total cost of the payment. At about $1.00, IFT per-transaction fees are about three times the per-transaction fees for regular Internet and mobile payments. The charge based on the transaction amount is the same across all three types of payments at approximately $0.07 per $1.00. Finally, the cap on the total price per payment is $5.00 for IFT payments, compared with $1.40 for regular Internet and mobile payments. It should be noted that IFT is differentiated from the pure RTGS wire transfers, not only in terms of operational process and timing (up to a one-hour delay for IFT compared with real-time for RTGS) but also in pricing. In the event that a bank client requests RTGS as the payment method, an even higher premium is charged.

Switzerland

Credit transfers have a long history in Switzerland, where the postal service has offered giro payments using a national standard format for over 100 years. (The credit transfer format known as *Einzahlungsschein* [credit slip] dates to 1906 and prevails to this day in a comparable form.) Traditionally, a credit slip has been used to initiate recurring and one-off payments, either over-the-counter at the post office or bank or, more recently, through the mail. The payee company would send a credit slip to the payer with pertinent information filled out, including bank/post and personal address; account number; and, if relevant, a reference number to assist the payee company in processing the payment. For payment purposes, account details are typically not perceived as confidential information by Swiss consumers and companies and are provided on a need-to-know basis to facilitate payments.

Today, IFT is available to businesses and individuals as an extension of the traditional credit slip. In addition to the traditional paper method, IFT is available through Internet banking and ATMs.[17] To illustrate the payer experience with IFT, imagine a computer terminal securely connected to a bank or PostFinance (the Swiss Post's financial institution) website. The payer clicks on "making payments" and receives a menu of choices among different types of credit slips, for example, payments to accounts at the same bank, at a different bank, payments with or without reference numbers, and so on. When it is selected, a digital credit slip opens and the payer fills out the necessary fields using the information received from the payee company. To reduce manual intervention, electronic payment-slip readers can be used. When the payer completes the instructions, the "electronic credit slip" is immediately verified by the system online and, assuming it is complete and correct, delivered to the bank for processing. The payer would typically not be aware of the particular infrastructure used to settle payments.

Credit transfers are typically settled through the Swiss RTGS system, called Swiss Interbank Clearing (SIC). This system is overseen by the Swiss National Bank (SNB) and operated by SIX Interbank Clearing Ltd. on behalf of the SNB. Swiss Interbank Clearing is owned by the Swiss commercial banks and PostFinance. General-purpose credit transfers have been more widely settled in SIC since PostFinance became a participant in 2001. The extension of SIC services beyond traditional large-value transfers is a cooperative development involving the commercial banks, PostFinance, and the central bank, and reflects their collective interest in supporting more efficient credit transfers, in this case making greater use of SIC and avoiding duplicative infrastructure for processing small-value payments. In this way, the banking system benefits from economies of scale in operations and pooling of liquidity. In addition, standards are followed to facilitate efficient processing (for example, increasing use of the international bank account number or IBAN) for routing information.

Pricing of IFT payments in Switzerland depends on the bank providing the service and the customer segment being served. Banks often include consumer payments as a component part of their bundled account service packages. Charges for account service packages depend on the balance that is maintained. Domestic payments would typically not carry a per-transaction charge. An exception would be paper payments that require manual processing steps for the banks or PostFinance. These payments would typically carry a surcharge as an incentive for the customer to use online banking.

United Kingdom

Faster Payments is a new IFT service in the UK that makes near-real-time and irrevocable credit transfers available to all bank customers at nonpremium prices. Introduced in May 2008, Faster Payments is available across the banking industry and is supported by common rules and a shared processing infrastructure. Faster Payments is a voluntary initiative of the banking industry, agreed to by the Payment System Task Force, which was organized and chaired by the UK's Office of Fair Trading (OFT). The OFT organized the task force in response to a mandate from the Chancellor of the Exchequer. The official mandate was reinforced by the threat of government-sponsored legislation to remedy perceived inefficiencies in the payment system, resulting from insufficient competition and overly slow cooperation among banks. Of principal concern to the government was a three-day delay in the interbank clearing of electronic payments.

The Payment System Task Force told the payments industry to devise a same-day service. The industry's response was to propose a near-real-time service, delivered through a special purpose infrastructure designed and operated by VocaLink. The company that is responsible for the Faster Payments Service (a name that is acquiring a brand identity for purposes of marketing the service to the public) is the CHAPS Clearing Company. The company provides two main services: CHAPS Sterling for systemically important payments and Faster Payments for time-dependent payments.

The 13 banks that originally agreed to develop the service now originate Faster Payments on behalf of their customers, and approximately 68 credit institutions, representing an estimated 90 percent of all transaction accounts in the UK, receive such payments. Membership in the Faster Payments Service is open to all credit institutions that have settlement accounts with the Bank of England and can connect their networks to the payment system infrastructure continuously, 24 hours a day, seven days a week. Indirect access is also permitted, whereby an institution offers the Faster Payments Service and settles through a member.

Customers can originate Faster Payments through their banks either by phone or Internet connection 24 hours a day, seven days a week; it is estimated that approximately two-thirds of all UK phone and Internet payments are now made by this method. Support for mobile Faster Payments is an important component of the UK's payment system strategy; it is seen by some as a viable alternative to reliance on the paper check (VocaLink and PriceWaterhouseCoopers, 2009). One-off payments are received by the beneficiary usually within minutes, but always within two hours. These one-off payments can be ordered on the payment date or submitted as forward-dated payments to be made on designated days in the future. Standing order payments are also possible, although these will be processed for same-day settlement and then only on bank working days. A direct corporate access feature has recently been added that enables companies with large volumes of payments to submit files directly to the Faster Payments Service infrastructure, provided they are sponsored by a member bank. This new feature is intended to increase the attractiveness of the service for firms that have a large number of expenses to pay, including payrolls, and is analogous to the services provided to corporate users of the ACH system in the UK.

A Faster Payment becomes final at the time the sending bank submits the transaction to the processing system; sending banks manage their risk by authenticating the instruction received from the originator of the payment and checking the customer's account to ensure that the balance is sufficient to fund the payment order. The Faster Payments Service processing system verifies that all of the required details are provided in the proper format and forwards the payment to the receiving bank. The receiving bank verifies that the funds are being directed to a valid account and sends a validation message back to the Faster Payments Service. The receiving bank is then credited with the funds. Confirmations of complete transactions are issued to the sender and receiver.

The prices charged for Faster Payments are a fraction of those charged for traditional CHAPS transfers, which can cost up to \$35.00 each. Marketing information published by banks indicates that per transaction prices are below \$1.25, ranging downward to about \$0.50. Transactions for retail customers are typically free. A size limit for transfers of GBP 100,000 has been set as a risk-management measure; this may be raised or eliminated in the future.

Summary

The four case studies are summarized in table 2. For each country, the table identifies the catalyst behind the introduction of the service, the delivery channels through which the banks provide the services to their customers, the back-end system for clearing and settling payments, the routing number scheme, and the prevailing fee structure. The four case studies illustrate two general approaches to interbank IFT processing. In Mexico and Switzerland, the national RTGS systems are relied upon for interbank processing, extending existing RTGS functionality to a broader set of underlying payments. In South Africa and the UK, the banks have created new, shared utilities that handle all of the interbank processing for the individual transactions and, in turn, rely on the national RTGS for final interbank settlement of netted IFT transfers periodically throughout the day.

In two of the four cases (Mexico and the UK), public authorities led in motivating a coordinated response across the banking system. In Mexico, the central bank served as catalyst and did so in part through its operational role as a provider of RTGS services. In the UK, the OFT, which shares responsibility for aspects of payment system oversight with the central bank, provided the motivation as a regulator concerned about the quality of payment services available to the general public. In contrast, in South Africa and Switzerland, banks identified an unmet service need (and opportunity) and took the lead, enlisting the central bank to provide support where necessary.

Table 2 highlights the areas where banks cooperate and compete in the provision of IFT services. Cooperation in planning is necessary to support nationwide services. In South Africa and the UK, the operational cooperation extends to governance over creation and enforcement of the rules that apply to the IFT network, as well as sharing in the investment and ongoing operating costs for the interbank processing system. With regard to routing of payments, note that only in Switzerland has the banking system adopted a standard routing number scheme, which facilitates processing for all parties to transactions and, further, makes it easier for senders and receivers of payments to manage the exchange of bank and account number information that is needed to route the transactions efficiently and accurately. As we describe later (in note 18), in the UK the banking clearinghouse provides bank routing information directly to the public.

The last column in the table summarizes the price structures and prices that apply to general-purpose IFT. In each case except South Africa, the price structure is essentially "cost-plus," that is, fees are based directly on the cost of production plus a markup reflecting service value and profit. In the case of South Africa, the banks not only charge per-transaction fees, but also an ad valorem fee component related to the

TABLE 2

International experience with immediate funds transfer

Country	Catalyst	Channel	Clearing and settlement	Routing	Fee
Mexico	Central bank	Online banking, mobile	RTGS (SPEI)	BAN[a]	Fixed per transaction (could be bundled), \$0.35–\$2.50
South Africa	Commercial banks	Online banking, mobile, over-the-counter	Real-Time Clearing (RTC)	BAN	Fixed per transaction, \$1.00 + ad valorem, \$0.07/\$1.00)
Switzerland	Majority of banks and central bank	Online banking, ATM, over-the-counter	RTGS (SIC)	BAN, IBAN[b]	Typically bundled with account service fees
UK (Faster Payments)	Competition authority	Online banking, mobile, direct corporate access	Faster Payments Service (FPS)	BAN	Fixed per transaction, typically free of explicit charges for retail customers, \$0.50–\$1.25

[a]Bank account number.
[b]International bank account number.

value of the transaction; this is similar to payment card price structures. The two approaches to pricing highlight an important two-part public policy question concerning the optimal way to price payment network services when credit risk is mitigated through the use of the immediate funds transfer model. First, is par clearing (receipt of the amount designated in the payment without deductions) a desirable goal? Second, can and should prices charged to end-users be based on production costs?

Issues with IFT implementation

What are the business and public policy issues that would need to be considered prior to the national introduction of IFT as a general-purpose means of payment in the U.S.? Three primary issues in addition to pricing are network reach, payment routing, and governance. Each of these issues has practical implications for the feasibility of IFT as a new payment service and each is important from a public policy perspective.

Network reach

IFT services are now available in the U.S., but are limited to closed proprietary networks. The process of clearing and settlement for these proprietary networks works efficiently only for the members who use a particular service provider's technology. In the case of a transfer destined for a receiver who is not a member of the proprietary network, the transaction must be routed through a bank payment system, such as ACH, using the national banking network. From a public policy perspective, the emergence of multiple, incompatible, and proprietary payment networks is not an efficient or effective way to provide IFT services.

A national clearing and settlement mechanism, however, does not guarantee that the payment network supporting an instrument such as IFT will connect all bank deposit accounts. As illustrated by the case studies, bankers may not be required to provide the service to their customers by regulation or by the terms of their clearinghouse memberships. An obvious practical problem with voluntary network participation, well illustrated in the case of Faster Payments in the UK, is that senders need to know whether their intended receivers hold accounts at a bank that can receive IFT transfers. A national directory sponsored by the UK clearinghouse is available online to help senders get this information as efficiently as possible.[18]

While not the subject of this article, the chartering and regulatory status of new, nonbank suppliers of payment services also has a bearing on the network reach issue. The innovators should not be prohibited from joining and helping stimulate improvements in the banking payment network by offering payment accounts, so long as they can meet basic tests of soundness and reliability, as do regular banks. As members of banking clearinghouses and associations, the nonbank innovators would contribute to the bank payment network's expansion. Moreover, to the extent that they innovate through the use of "disruptive technologies," these nonbank companies would stimulate technological innovation in services such as IFT. The U.S. financial regulatory authorities should consider how payment innovation can be encouraged by allowing nonbank firms to offer deposit accounts on terms that are reasonable and prudent.[19]

Payment routing

The principal operational advantage of payments such as checks and electronic direct debits is that they provide routing information that the payer would otherwise have to request. On a paper check, for example, the payer's bank routing number and account number are printed in magnetic ink at the bottom of the check. Thus, the payment instruction automatically contains the data needed by the payee's bank to present the instrument for payment. Routing information is provided with debit card payment instructions as well. For electronic credits and IFT, the payer needs to obtain payee routing information and provide that information to its bank. Acquisition of this information adds complexity and cost, especially for transactions between two parties that are not well known to one another.

Account numbers are sometimes considered to be part of one's "transactional identity," which is sensitive information that should be protected. Because of this concern, receivers may be reluctant to give their account number to a payer for an IFT payment. Such concerns, however, should be reduced by the IFT payment flow and authorization model. First, IFT results in money deposited to the receiver's account, not withdrawn from it. Second, bank controls are designed to restrict the power to initiate transfers of funds to properly authenticated parties. Thus, there is limited opportunity for anyone to fraudulently order an IFT based on knowledge of an account number and routing number.

As mentioned, paper checks contain complete routing information that is in plain view to anyone handling the check. This is prima facie evidence that routing information is not unduly sensitive. It is not considered so in the countries examined in connection with the four case studies. Further, it is notable that the IFT payment services provided by nonbanks often rely on widely known and used "addresses" for routing and information exchange over networks, including telephone numbers and email addresses. The new approaches to routing appear to point to the serviceability of highly public addresses for transferring financial information, including funds transfers, in a well-controlled environment with strong information security protections.

A somewhat broader issue that arises when considering routing of payments and the use of account numbers is that of standardization and portability of financial addresses. If bank account numbers are not standardized across the banking system and are not portable, bank numbers change whenever an account holder changes banks. Switching banks becomes more complex because all established payment relationships must be updated with new account information. Progressive banking practice and good public policy call for both standardization and portability of bank account numbers, both to increase the efficiency of the payment system and to increase competition among banks by making it harder to lock in customer relationships through high switching costs. This is not an unreasonable expectation in an information-intensive industry like banking. Public policy that is concerned with the efficiency and competitiveness of payment services could be informed by practices and expectations in other information-intensive industries, for example, telecommunications.[20]

Payment system governance

Each of the four case studies discussed in this article provides an example of payment system innovation coordinated at the national level. The catalyst may be from the public sector (central bank or other governmental authority, such as the UK's Office of Fair Trading) or the private sector (groups of banks), but in each case IFT innovation proved successful due to a national governance approach. In addition, the governance approach followed in the four countries recognizes the boundary between cooperation and competition among banks.

This type of national, coordinated approach would be difficult to achieve in the U.S. in light of its highly decentralized payment system management, which is reflected in part by the absence of a truly national clearinghouse. Currently, multiple publicly and privately operated payment systems operate in parallel in a competitive environment. Sweeping national change in the U.S. payment system in this century so far has come about through legislation—the 2003 passage of the Check Clearing for the 21st Century Act, which facilitated electronic check clearing; and the 2010 Wall Street Reform and Consumer Protection Act, which mandated limits on fees that banks charge merchants for debit card transactions. Without an explicit legislative mandate or some other form of encouragement from the government, it is unlikely that banks in the U.S. will find a cooperative basis for IFT innovation. In addition, because IFT may disrupt banks' revenues from high-priced wire transfer services, coordination and cooperation may not be readily forthcoming. Further, unless IFT clearing and settlement relies on existing mechanisms (as in the cases of Mexico and Switzerland), a national IFT system may have high start-up costs that the industry might be unwilling to bear. Overall, the complexity involved with implementing a national IFT solution may be unworkable within the existing U.S. banking structure.

Conclusion

General-purpose IFT is a means of payment that offers attractive combinations of attributes to both senders and receivers, such as certainty, speed, control, and versatility, all at relatively low cost. There is evidence of strong latent demand for IFT in the U.S. by individuals, businesses, and governments, but to date this demand is being met only to a limited extent and principally by nonbank providers of payment services. To satisfy the demand for IFT, it will be necessary to provide access to money held in banks by linking all bank deposit accounts through an immediate if not real-time clearing and settlement system.

Within the last few years, IFT has become a fully functional nationwide means of payment in a number of countries, including four that we have examined in detail in this article. International experience with IFT shows that technology is a necessary but not sufficient condition for innovation in payments and that enabling real-time and universal access to deposit accounts at banks is the key to meeting the public's needs for more certain, faster, and universal payment services. Perhaps the most critical enabling factor is strong sponsorship by a national body with the responsibility and motivation to stimulate continuous improvement in the national payment system. This body might be a consortium of private banks collaborating through a national payment association, a public authority such as the central bank, or a public–private partnership. It is not clear that such sponsorship can be readily found in the U.S., at least not at the present time, because there is no national body that takes responsibility for the development of the national payment system. As a consequence, IFT and other national payment innovations are likely to progress in a halting and incomplete manner and at a pace that lags innovation that is observable in other countries, such as those examined in this article.

NOTES

[1]See https://www.paypal.com.

[2]The exception to the norm is Japan, where cash is more widely used than in any other industrialized country due to factors such as relatively low crime rates, effective anti-counterfeiting measures, and low-cost nationwide ATM networks (BIS, 2003).

[3]Depository institutions include banks, thrifts, and credit unions. In this article, the term "bank" means all depository institutions.

[4]A full discussion of credit transfers and debit transfers is provided in the appendix.

[5]Two-sided markets require the participation of two separate parties in order to succeed (Rochet and Tirole, 2003). A sender and receiver of a payment must use the same payment system in order to exchange monetary value.

[6]Some checks are converted to electronic format at the point of acceptance and are cleared through the automated clearinghouse (ACH) network, as described later.

[7]A cash payment is also a credit transfer.

[8]As noted in the introduction, the most prominent example of the first approach is PayPal. Examples of the second approach include CashEdge (www.cashedge.com/) and Obopay (https://www.obopay.com/consumer/welcome.shtml).

[9]These closed proprietary networks were first described by Kuttner and McAndrews (2001).

[10]The same description applies to transfers among accounts held at the same bank, called intrabank or "on us" transfers.

[11]Some cardholders are aware of the delay in the transfer of deposit money and "play the float" with these transactions. For those cardholders, debit card transactions are not perceived as immediate.

[12]Debit card cost structure has become controversial to the point that recent banking reform legislation directs the Board of Governors of the Federal Reserve System to regulate merchant fees and includes a provision to allow merchants to offer discounts for customers who pay with cash or check. (Wall Street Reform and Consumer Protection Act, §1075).

[13]Respondents included corporations, technology firms, banks, payment processors, and infrastructure providers.

[14]In a joint April 26, 2010, press release, the Federal Reserve System and The Clearing House Payments Company L.L.C. announced plans to implement enhanced message formats to support extended-character business remittance information for U.S. dollar wire transfers on November 11, 2011.

[15]The findings in this section are based on correspondence with central bankers and examination of the public websites of payment services providers, including commercial and central banks and the financial services arm of the post office. The authors acknowledge and are grateful for the assistance provided by Ricardo Medina (Banco de México), Dave Mitchell and Mike Stocks (South African Reserve Bank), Philipp Haene and Dave Maurer (Swiss National Bank), and Paul Smee (UK Payments Council), none of whom bear any responsibility for the descriptions, analysis, and conclusions presented in this article.

[16]Straight-through processing (STP) is an operational design based on standards that allow for fully automated processing of a payment from its origination by the payer to its receipt by the payee.

[17]Also, mobile payments for small accounts using cell phones have been introduced by PostFinance for payments between PostFinance account holders.

[18]The directory can be found at www.ukpayments.org.uk/sort_code_checker/.

[19]One approach would be to charter so-called "narrow banks" that specialize in payments. This approach has the advantage of encouraging innovation, while at the same time prudently extending the public safety net of deposit insurance to new market entrants (Litan, 1987).

[20]Mobile phone numbers, for example, are portable from one carrier to another.

APPENDIX: MODELS OF PAYMENT TRANSACTIONS

Two basic payment models frame the classification of all types of payment transactions. These are 1) credit transfers and 2) debit transfers. The end result of these transfers is the same: Deposit money is transferred from payer to payee. The process that results in the transfer of deposit money, however, is quite different. In a credit transfer, deposit money is moved directly from a payer's or sender's transaction account to a payee's or receiver's account. A credit transfer is sometimes referred to as a "credit push" payment, meaning that money is delivered directly to the receiver based on instructions made by the sender to the sender's bank. In a debit transfer, deposit money is moved in a less direct manner and requires the receiver to request a transfer from the sender's bank, based on authorizing instructions provided by the sender. A debit transfer is sometimes referred to as a "debit pull" payment, meaning that the receiver must present the sender's instruction to the sender's bank before deposit money is transferred.

Operationally, payment transactions are more complex than described in the foregoing paragraph. For purposes of modeling, a generic payment transaction can be visualized as consisting of two discrete information flows involving "instructions" and "funds movement," which are illustrated in figures A1 and A2 for credit and debit payments, respectively.[1] Instructions are shown as solid lines and funds movements are shown as dotted lines.

For credit transfers, as shown in figure A1, a sender instructs his/her bank to deliver funds to a designated receiver.[2] These instructions result in a debit to the sender's transaction account and initiate movement of funds from the sender's bank to the receiver's bank and credit to the receiver's account. For debit transfers, as shown in figure A2, a sender does not directly instruct his/her bank to transfer funds. Instead, payment instructions follow a chain from sender to receiver, then from the receiver to his/her bank, and finally from the receiver's bank to the sender's bank to transfer money from the sender's account.[3] These instructions result in a credit to the receiver's account; however, because the receiver's bank is uncertain at the time instructions are delivered to the sender's bank whether the sender's bank will honor the instructions, final credit to the receiver's account is delayed by the time it takes the sending bank to determine whether it will honor the payment. Accordingly, funds transferred by the debit transfer method are typically made available to receivers as provisional funds and are subject to reversal. If the sender's bank honors the instructions, then the sender's account is debited and provisional funds become final.

[1]Depending on the payment method and the system used, funds movement may also include data related to the payment, such as invoice or remittance information and reference numbers.

[2]The discussion in this paragraph closely follows Geva (2009).

[3]For both credit and debit transfers, one or more intermediary banks may stand between a sender's bank and a receiver's bank to execute the transfer of deposit money. In addition, senders in both models may use agents, such as a payroll processing company, to initiate instructions on their behalf.

FIGURE A1

Credit transfer

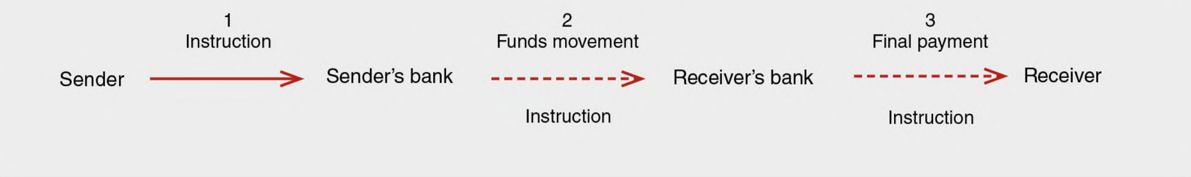
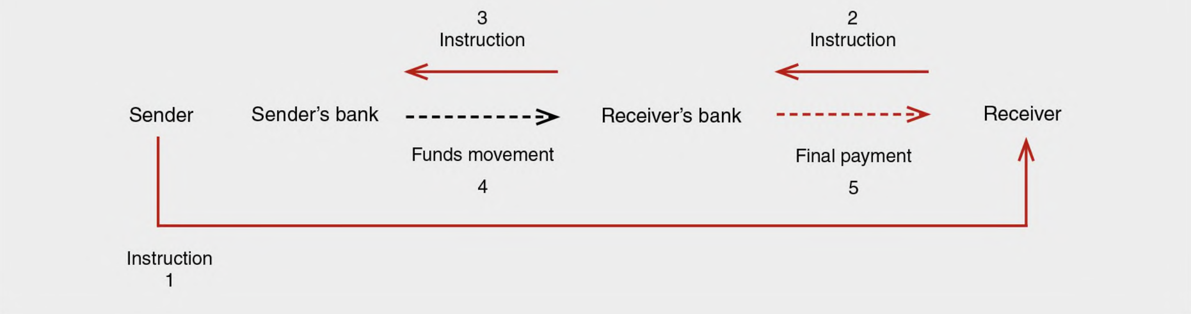


FIGURE A2

Debit transfer



REFERENCES

Allsopp, P., B. Summers, and J. Veale, 2009, "The evolution of real-time gross settlement: Access, liquidity and credit, and pricing," Financial Infrastructure Series: Payment Systems Policy and Research, The World Bank, February.

Association for Financial Professionals, 2010, "2010 AFP electronic payments: Report of survey results," November.

Bank for International Settlements, 2003, *Payment and Settlement Systems in Selected Countries*, Committee on Payment and Settlement Systems, CPSS Publications No. 53, Basel, Switzerland, April.

______________, 2001, "Core principles for systemically important payment systems," report, Committee on Payment and Settlement Systems, CPSS Publications No. 43, Basel, Switzerland, January.

Board of Governors of the Federal Reserve System, 2010, *The 2010 Federal Reserve Payments Study—Noncash Payment Trends in the United States: 2006–2009*, research report, December 8.

______________, 2007, *The 2007 Federal Reserve Payments Study—Noncash Payment Trends in the United States: 2003–2006*, research report, December 10.

______________, 2006, "A summary of the roundtable discussion of the role of wire transfer in making low-value payments," Payments System Policy Advisory Committee, report, Washington, DC, May, available at www.federalreserve.gov/paymentsystems/lowvaluepay/lowvaluepayments.pdf.

______________, 2004, *The 2004 Federal Reserve Payments Study—Analysis of Noncash Payments Trends in the United States: 2000–2003*, research report, December 15.

______________, 2002, "The future of retail electronic payments systems: Industry interviews and analysis," staff study, No. 175, December.

Boland Hill Media, LLC, 2010, "Latest ACH stats reflect consumer shift to electronic bill pay," *Digital Transactions News*, August 20, available at www.digitaltransactions.net/newsstory.cfm?newsid=2616.

Foster, K., E. Meijer, S. Schuh, and M. Zabek, 2010, "The 2008 survey of consumer payment choice," Public Policy Discussion Papers, Federal Reserve Bank of Boston, No. 09-10, April.

Gerdes, G.R., 2008, "Recent payment trends in the United States," *Federal Reserve Bulletin*, pp. A75–A106, October.

Geva, B., 2008, "Payment finality and discharge in funds transfers," *Chicago-Kent Law Review*, Vol. 83, No. 2, pp. 633–675.

Herbst-Murphy, S., 2010, "Trends and preferences in consumer payments: Lessons from the VISA payment panel study," Federal Reserve Bank of Philadelphia, discussion paper, May.

Hough, Daniel, George Simotas, Kim Allin, and John Hansen, 2010, "Shifting to debit," *BAI Banking Strategies*, Bank Administration Institute, October 20.

Kuttner, K., and J. McAndrews, 2001, "Personal on-line payments," *Economic Policy Review*, Federal Reserve Bank of New York, Vol. 7, No. 3, December, pp. 35–50.

Litan, R., 1987, *What Should Banks Do?*, Washington, DC: The Brookings Institution.

McKinsey & Company, 2009, "2009 U.S. payment revenues," unpublished PowerPoint slide.

Rochet, J-C., and J. Tirole, 2003, "Platform competition in two-sided markets," *Journal of the European Economics Association*, Vol. 1, No. 4, June, pp. 990–1029.

Rysman, M., 2010, "Consumer payment choice: Measurement topics," in *The Changing Retail Payments Landscape: What Role for Central Banks?*, Federal Reserve Bank of Kansas City, pp. 61–81.

Shevlin, R., J. Fishman, and G. Bezard, 2010, "Sizing person-to-person payments in the United States, United Kingdom, and Australia," Aite Group, LLC, report, No. 2010110301, October.

Smith, K., 2010, "Consumer payment choice: Measurement topics—Commentary," in *The Changing Retail Payments Landscape: What Role for Central Banks?*, Federal Reserve Bank of Kansas City, pp. 83–92.

VocaLink and PriceWaterhouseCoopers, 2009, "Tomorrow happened yesterday: How banks are building a business case for Faster Payments," report, available at www.finextra.com/finextra-downloads/newsdocs/tomorrow-happened-yesterday---how-banks-are-building-a-business-case-for-faster-payments.pdf.

Wells, Kirstin E., 1996, "Are checks overused?," *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 20, No. 4, Fall, pp. 2–12.

World Bank Group, The, 2008, *Payments Systems Worldwide: A Snapshot. Outcomes of the Global Payment Systems Survey 2008*, Financial Infrastructure Policy and Research Series.

How do benefit adjustments for government transfer programs compare with their participants' inflation experiences?

Leslie McGranahan and Anna L. Paulson

Introduction and summary

Millions of Americans rely on government transfer programs as a way to make ends meet during a temporary setback or as their main source of income during retirement. Whether individuals qualify for unemployment assistance, Temporary Assistance for Needy Families (TANF), Social Security,[1] or Supplemental Security Income (SSI), the level of benefits they will receive is affected by how the benefits are adjusted to deal with inflation—the general increase in prices for goods and services over time. Changes in benefit levels to address inflation—that is, cost-of-living adjustments (COLAs)—are determined by formulas that vary depending on the program in question. Adjustments to some programs' benefits are made automatically based on a government inflation index, while adjustments to others require legislative action.

COLAs can have a substantial impact on the welfare of transfer program participants. Those who receive benefits from a program for a long time are particularly affected by the formulas determining COLAs. In addition, COLAs can have a large impact on transfer program costs. For example, the bipartisan National Commission on Fiscal Responsibility and Reform (chaired by Democrat Erskine Bowles and Republican Alan Simpson) recently proposed changing the way COLAs are made for Social Security benefits. By making Social Security COLAs using a chain-weighted Consumer Price Index (C-CPI),[2] as opposed to the current method that relies on the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W), benefits are expected to increase by about 0.3 percentage points less each year. This small change in the formula for determining the Social Security COLAs would significantly affect both the benefits received from the program and the program's costs. According to our calculations, if inflation measured by the CPI-W averaged 2.5 percent per year for the next 15 years, an individual receiving $25,000 in Social Security payments this year would receive 15 years from now an annual payment of $36,207. Under the chain-weighted formula, assuming inflation averaged 2.2 percent per year (0.3 percentage points less than under the current formula), the same individual would receive 15 years from now an annual payment of $34,650. According to the National Commission on Fiscal Responsibility and Reform (2010, pp. 54, 65), if the proposed change in the method for calculating COLAs for Social Security were enacted, this change would lead to savings of $89 billion over the period 2012–20 and would reduce the Social Security shortfall by 26 percent over 75 years.

Major U.S. transfer programs target individuals with particular characteristics—for example, single mothers and the elderly. These individuals are likely to have different spending patterns than the average individual. However, programmatic COLAs are typically based on aggregate inflation measures. Since different groups of individuals purchase different goods and services, they may face a rise in their cost of living that differs from that of the average household. For example, the elderly spend more on health care than the general population, and commuters spend more on transportation. If health care costs increase more rapidly than aggregate prices, then the inflation experienced by the elderly will be greater than general inflation. Similarly, if gas costs and therefore the costs of transportation increase rapidly, then commuters will face inflation that is higher than that of the general

Leslie McGranahan is a senior economist and Anna L. Paulson is a vice president and director of financial research in the Economic Research Department at the Federal Reserve Bank of Chicago. *The authors thank Bruce Meyer for helpful discussions and Daniel DiFranco, Lori Timmins, and Nathan Marwell for excellent research assistance.*

population. COLAs for major transfer programs do not typically account for these differences in spending patterns.

In this article, we measure the inflation experienced by different groups of people. We focus on groups that are likely recipients of federal benefits: the elderly, single mothers, individuals with less than a high school diploma, the disabled, and the poor. We compute group-specific inflation measures for the period 1980–2010, using data on group spending patterns from the U.S. Bureau of Labor Statistics' (BLS) *Consumer Expenditure Survey* in combination with item-specific inflation measures, also from the BLS. We then compare our group-specific inflation measures with the COLAs used for major transfer programs to evaluate whether program benefits that are adjusted using aggregate measures of inflation or using other means "keep up" with the inflation experienced by a specific group.

COLAs can affect the welfare of transfer program recipients in (at least) two ways: by determining the initial level of benefits that they receive and by determining how benefit payments grow during program participation. The latter is particularly important for programs like Social Security that individuals often participate in for a long time, and the former is a key factor of programs like TANF that individuals usually participate in for shorter periods.

We find that the elderly and, to a lesser extent, the disabled, the poor, and those with less than a high school diploma experienced higher inflation than the aggregate population from 1980 through 2010. Because the Social Security/SSI COLA is based on aggregate inflation, Social Security/SSI COLAs have been less than the price increases experienced by the elderly in most periods since 1980. More specifically, in 2010, Social Security benefits for an individual who had been on the program since 1980 would be 265 percent of their nominal 1980 value, while the cost of the items purchased by the average elderly household was 270 percent of their nominal 1980 value. Inflation faced by the disabled, while above aggregate inflation, has been slightly below the Social Security/SSI COLA because of nuances in COLA determination. Single mothers experienced lower inflation than the overall population during this period, but the inflation they faced was larger than the increases in benefits from the Aid to Families with Dependent Children (AFDC) program and TANF. Increases in welfare benefits from the AFDC and its successor program, TANF, in most states have been substantially below both aggregate inflation and the price increases faced by single mothers over the period 1980–2010. In addition, we find that the growth in benefits from the Supplemental Nutrition Assistance Program, or SNAP (formerly called the Food Stamp Program), has exceeded the inflation faced by single mothers, the disabled, the poor, and those with less than a high school diploma over the period 1980–2010, largely because of increases in benefit levels enacted as part of the American Recovery and Reinvestment Act (ARRA) of 2009.

During the recent recession and subsequent recovery, U.S. inflation has been atypically low. Also, during this period, there have been somewhat unusual COLAs for both Social Security/SSI benefits and SNAP benefits. Because this period is unique from both an inflation perspective and policy perspective, we break our analysis into two periods: 1980–2008 and 2008–10.

The rest of our article is organized as follows. In the next section, we discuss major U.S. transfer programs and report how benefits from these programs are adjusted for inflation. Then, we describe the characteristics of program recipients and compare them with the overall U.S. population. These comparisons are used to identify the groups whose inflation experiences we would like to investigate. Next, we compare the inflation experiences of these groups with the inflation experience of the aggregate U.S. population, and discuss how these comparisons were generated. We also compare group inflation experiences with programmatic COLAs. We highlight four programs in our analysis of COLAs: Social Security, SSI, TANF, and SNAP.[3] Finally, we review our conclusions and briefly discuss them in the context of the policy debate concerning COLAs, which has chiefly revolved around the Social Security program.

COLAs for major transfer programs

The federal government transfers money to many different recipient populations through a large variety of targeted programs. Table 1 lists all of the federal government's transfer programs with total direct payments and indirect payments (which are largely payments made via states) to individuals that exceeded $10 billion in fiscal year 2010, according to the 2012 federal budget. This table also lists the outlays on the program, the number of beneficiaries served, the way in which benefits or expenditures are adjusted for changes in the price level, and a brief description of the eligibility criteria. There are 18 such programs, which served a total of 379 million recipients in 2010, indicating that the average American is served by more than one of these programs.

Combined, these programs cost the federal government $2.2 trillion in fiscal year 2010 and made up over 95 percent of all federal payments to individuals. These programmatic expenses represented approximately 60 percent of all federal government outlays in fiscal

TABLE 1

Government transfer programs with payments exceeding $10 billion, fiscal year 2010

Program	Outlays for payments to individuals (in millions of dollars)	Beneficiaries (annual average in thousands)	Inflation adjustment/ benefit determination	Eligibility
Child nutrition programs (not including Women, Infants, and Children program and Commodity Supplemental Food Program) and Special Milk Program	16,430	34,889	Beneficiaries receive low-cost or free nutritionally balanced breakfasts and lunches. Payments to a state depend on changes in the Food Away from Home series in the CPI-U.	Children in school or child care from families with incomes at or below 130 percent of the FPL are eligible for free meals; those from families with incomes between 130 percent and 185 percent of the FPL are eligible for reduced price meals.
Civil Service Retirement System	69,407	2,523	CPI-W (July–Sept.) determines cost-of-living adjustment.	Retired federal government workers.
Earned income tax credit (refundable portion)[a]	54,712	21,743	Thresholds and maximum credit are adjusted by CPI-U (for the 12-month period ending August 31).	Low-income working individuals and families. For single-parent families with two children, annual income must be below $40,363; for single individuals, annual income must be below $13,460 (2010 tax year).
Hospital and medical care for veterans[b]	38,216	5,639	In kind. No maximum benefit level.	Individuals who actively served in the military (priority to those with service-connected disabilities and low income).
Housing assistance[c]	49,959	3,200	In kind. For public housing rental payment equal to 30 percent of monthly adjusted income.	Income below limits that are area specific (30 percent, 50 percent, and 80 percent of median).
Tenant-based rental assistance/ housing choice vouchers (Section 8)[c]	17,987	2,100	Benefit amount equal to the difference between 30 percent of adjusted household income and the public-housing-authority-determined payment standard (90 percent to 110 percent of the fair market rent, or FMR). FMR is set by the federal government as average gross rents (including utilities) for medium-quality apartments.	Income below some percentage (between 50 percent and 80 percent) of local area median. Housing authorities can have additional criteria. At least 75 percent of new households must have income at or below 30 percent of the area median.
Medicaid	272,771	59,339	In kind. No maximum benefit level.	States are required to serve pregnant women and children under six years old with income below 133 percent of the FPL, children aged six to 19 up to 100 percent of the FPL, and Supplemental Security Income recipients. States can include other groups.
Medicare: Hospital Insurance (Part A)	250,223	46,906	In kind. No premium for most beneficiaries.	Free if aged 65 years or older and individual or spouse is receiving Social Security or if under 65 and receiving Social Security disability benefits. Nonqualifying individuals aged 65 or older can pay for Part A coverage.

TABLE 1 (CONTINUED)

Government transfer programs with payments exceeding $10 billion, fiscal year 2010

Program	Outlays for payments to individuals (in millions of dollars)	Beneficiaries (annual average in thousands)	Inflation adjustment/ benefit determination	Eligibility
Medicare: Supplemental Medical Insurance (Part B)	268,945	43,569	Monthly premiums, deductibles, and co-insurance amounts are adjusted by the federal government, as determined by formula. Premium is required to be the amount needed to cover 25 percent of estimated program costs for enrollees aged 65 years and older.	Anyone eligible for Medicare Part A can pay a premium and enroll in Part B.
Military retirement	50,585	2,212	CPI-W (July–Sept.) determines cost-of-living adjustment.	Retired members of the military (no specific age requirement).
Refundable (additional) child credit[d]	22,659	18,160	Child credit = $1,000 nominal; no automatic adjustment. Set legislatively. Has been increased on occasion—Economic Growth and Tax Relief Reconciliation Act of 2001 doubled the credit. It will return to $500 in 2013.	Tax filers with children under 17 years old whose tax liability is not large enough to fully discharge the $1,000 per child credit (or the amount remaining after phaseouts) can get the additional child credit. Child credit is reduced by 5 percent of adjusted gross income (AGI) for married filers with AGI over $110,000 and single parents with AGI over $75,000. The maximum additional child tax credit is limited to 15 percent of earnings above a threshold ($3,000 in 2010; historically this threshold was indexed for inflation using CPI-U).
Social Security: Disability Insurance	123,507	9,822	CPI-W (July–Sept.) determines automatic cost-of-living adjustment.	Individuals who worked in Social Security-covered employment for a sufficient amount of time, are under retirement age, and are unable to perform any substantial work for at least one year.
Social Security: Old-Age and Survivors Insurance	576,578	43,110	CPI-W (July–Sept.) determines automatic cost-of-living adjustment.	Individuals who contributed to the program for 40 quarters or more and who have reached the minimum retirement age (62 years old). Surviving spouses and children of contributors are also eligible.
Student assistance—U.S. Department of Education and other (primarily Pell Grants and subsidized Stafford Loans)	46,768	20,638	Maximum loan amounts are nominally fixed, but periodically adjusted.	Pell Grants for students with family annual income below $45,000 (most with family annual income below $20,000). Subsidized Stafford Loans based on "financial need."
Supplemental Nutrition Assistance Program (officially known as the Food Stamp Program until October 1, 2008, and still commonly referred to by this name)	70,492	40,302	Changes in the cost of items (using the Consumer Price Index) in a "market basket" based on the Thrifty Food Plan (TFP) from June determine benefit levels starting in October. Revised TFP originally relied on prices paid by low-income households. The American Recovery and Reinvestment Act of 2009 led to unusual adjustments.	Household gross monthly income at or below 130 percent of the FPL; net income (net of allowances) below 100 percent of the FPL; and assets below $2,000 (or below $3,000 if one person is aged 60 years or older or is disabled).

TABLE 1 (CONTINUED)

Government transfer programs with payments exceeding $10 billion, fiscal year 2010

Program	Outlays for payments to individuals (in millions of dollars)	Beneficiaries (annual average in thousands)	Inflation adjustment/ benefit determination	Eligibility
Supplemental Security Income (SSI)	43,886	7,522	CPI-W (July–Sept.) determines automatic cost-of-living adjustment.	Aged (65 years old or older), blind, or disabled with assets below $2,000 per individual or $3,000 per couple. Benefit phases out with income. Benefit fully phased out for those with monthly income (net of allowances) greater than $694 per individual and $1,031 per couple.
Temporary Assistance for Needy Families[e]	21,936	4,594	Maximum benefits have been fixed at nominal levels since 1996 in many jurisdictions. Legislated changes in others. No inflation adjustment in grants to states.	Families with dependent children and pregnant women with income and assets below a state-determined threshold. Phases out at about 75 percent of FPL in average state. Households subject to work requirements and time limits.
Unemployment assistance	158,263	11,429	Benefits set at a fraction of weekly wage up to a maximum. In 36 jurisdictions, maximum benefit levels automatically adjust according to weekly wages of covered employees. In remainder, maximum benefits are set at a fixed dollar amount that can be changed by legislation.	Recently unemployed workers who are unemployed through no fault of their own, earned qualifying wages, and are actively seeking work.
Veterans service-connected compensation	43,377	3,498	CPI-W (July–Sept.) determines cost-of-living adjustment. Not automatic. Needs legislative approval. Typically granted unanimously.	Veterans who are at least 10 percent disabled as a result of military service, and their survivors.

[a]The number of earned income tax credit (EITC) beneficiaries is based on the number of tax returns with refundable EITC for tax year 2008 (see Internal Revenue Service, Statistics of Income, 2010, table 2.5).

[b]The number of beneficiaries of hospital and medical care for veterans is based on the total number of patients from www.va.gov/vetdata/Utilization.asp.

[c]Housing assistance and tenant-based rental assistance beneficiary estimates are from the U.S. Department of Housing and Urban Development (2011). Tenant-based rental assistance/housing choice vouchers (Section 8) is a subcategory of housing assistance.

[d]The number of refundable (additional) child credit beneficiaries is based on number of tax returns with additional child tax credit for tax year 2008 (see Internal Revenue Service, Statistics of Income, 2010, table A).

[e]The Temporary Assistance for Needy Families caseload data for 2010 are from www.acf.hhs.gov/programs/ofa/data-reports/caseload/caseload2010.htm. This program became the successor to the Aid to Families with Dependent Children program in 1997.

Notes: CPI-U means Consumer Price Index for All Urban Consumers. CPI-W means Consumer Price Index for Urban Wage Earners and Clerical Workers. FPL means federal poverty line. See note 6 for the definition of market basket. In some cases, there are minor differences between the coverage of the cost and beneficiary numbers.

Sources: Outlays data from White House, Office of Management and Budget (2011b), table 11.3; beneficiaries data for most programs from White House, Office of Management and Budget (2011a), table 27-5; and some beneficiaries data, inflation adjustment/benefit determination data, and eligibility data from various government sources.

year 2010.[4] For some of these programs, in particular Medicaid and unemployment assistance, state governments also expend significant sums of money. The only state dollars that are included in table 1 are those that were funded by transfers from the federal government.

Inflation adjustment/benefit determination information is presented in table 1 (pp. 115–117). The inflation adjustment/benefit determination column explains in detail how program benefits are adjusted for inflation for an individual once he is already enrolled in the program. For many programs such as TANF and SNAP, the level of benefits upon initial enrollment is set using the same formula. For other programs, initial benefits are set using a different formula. For example, the initial Social Security benefit level depends on earnings over the recipient's working life.

The inflation adjustment/benefit determination column in table 1 shows that there are four main types of adjustments used by these programs. First, some programs adjust benefit levels based on an aggregate inflation index—either the Consumer Price Index for All Urban Consumers (CPI-U) or CPI-W. The programs in this category are Civil Service Retirement System, earned income tax credit (EITC), military retirement, Social Security (both Old-Age and Survivors Insurance and Disability Insurance), SSI, and veterans service-connected compensation.[5] These tend to be the large income-transfer programs. The CPI-U and CPI-W are the two aggregate indexes released by the BLS. They are both consumer price indexes and as such represent changes in the cost of "market baskets"[6] consumed by different demographic groups. The CPI-W is calculated based on price increases for goods consumed by households for which at least half of household income comes from the earnings of workers in hourly wage or clerical jobs. This index represents about 32 percent of the U.S. population. The CPI-U is based on the market basket of all urban consumers; it represents 87 percent of the population. Second, some programs have benefits that are linked to price growth in a particular category. The programs in this category are child nutrition programs (in particular, the National School Lunch and Breakfast Programs) and the Special Milk Program; SNAP; and tenant-based rental assistance (Section 8 housing assistance). These programs are supporting consumption in a specific category, and therefore, their benefits are linked to price growth in that category. Third, some programs have no inflation adjustment because benefits are in kind. These programs are hospital and medical care for veterans, Medicare (Parts A and B), Medicaid, and non-Section 8 housing assistance. While there is no explicit benefit adjustment, the value of the benefits increases as the cost of the underlying good increases. A final set of programs has benefits that are nominally fixed in value. Benefit amounts can be changed through legislation. These programs are student assistance, the refundable (additional) child tax credit, and TANF. Unemployment assistance, for which the increases in benefits are based on wage growth, does not fall into any of these categories. No programs are linked to the broad-based expenditure needs of the program's recipient population.

If we combine the program costs for the programs that link to the CPI-U and CPI-W, we find that about $960 billion in annual expenditures was linked to these indexes, representing about 28 percent of total federal expenditures for 2010.

In addition to indexing benefit levels to inflation, the federal government indexes eligibility criteria for many transfer programs to inflation. In many cases eligibility is based on federal poverty guidelines, which are indexed to the CPI-U. Also, many features of the tax code, such as personal exemptions and tax brackets, are indexed (Hanson and Andrews, 2008).[7]

Characteristics of program participants

Next, we are interested in finding out what percentage of the population benefits from these programs and which demographic groups are especially dependent on benefit payments. Benefit levels, and hence COLAs, are especially important to households that receive a large fraction of their income from federal government transfer programs. We divide the population in six different ways—by education, age, disability status, family structure, veteran status, and poverty status. We choose these six methods of segmenting the population because they are in keeping with program eligibility standards and because the groups based on these different division criteria are some of the groups that are highlighted in other research on transfer program participation (see, for instance, Meyer and Rosenbaum, 2001, and Haveman et al., 2003). In addition, we are interested in groups whose recipient status tends to be fairly persistent. Gaps between programmatic inflation adjustments and household expenditure growth will be more relevant if households benefit from programs over long periods so that the gaps are compounded over time. Because of this issue, we do not look at population groups based on work status because employment status has historically been fluid.

In box 1, we describe the criteria we use for the inclusion of households in the groups listed there. As delineated in box 1, our definition of the disabled only includes those individuals who do not have a job rather than all individuals with a disability. In table 2, we present results showing what fraction of households

BOX 1

Demographic group variable descriptions

Variable	Description
Less than high school diploma	Neither the reference person nor spouse completed high school.
High school graduate, no college	Reference person or spouse obtained a high school diploma; neither the reference person nor spouse pursued education beyond the high school level.
Some college or more	Reference person or spouse pursued education beyond the high school level.
Elderly	Reference person or spouse is at least 65 years old.
Disabled	Reference person or spouse is out of work because of a chronic health condition or disability.
Single mother	Reference person is an unmarried female aged 18–64 years old; the reference person's child who is younger than 18 years old lives in the household.
Other parent	Reference person aged 18–64 years old is either an unmarried male or a married male or female; the reference person's child who is younger than 18 years old lives in the household.
Nonparent	Reference person is aged 18–64 years old; the reference person has no children who are younger than 18 years old living with him or her in the household.
Veteran	Reference person or spouse served on active duty in the U.S. Armed Forces at some point in his or her lifetime (currently active members of the Armed Forces are included).
Poor	Household's income was below the poverty line (adjusted for household size and composition) during the last month of reference period.

in these groups were recipients of benefits from the different programs listed in table 1 (pp. 115–117). This information is calculated from wave 4 of the 2004 panel of the U.S. Census Bureau's *Survey of Income and Program Participation* (SIPP), corresponding to the January–April period of 2005. The results displayed in table 2 are consistent with the eligibility criteria outlined in table 1. For example, 49 percent of families headed by a single mother receive a benefit from the National School Lunch and Breakfast Programs, while only 2 percent of households without children report receiving a benefit from these programs. Similarly, the vast majority of the elderly households receive both Medicare and Social Security.

In table 3 (p. 122), we show the median and average numbers of transfer programs that members of these different groups participated in. The median household of the overall sample receives a benefit from one of these programs (table 3, final row). For many groups, the median household participates in no benefit programs—these groups are the households with some college or more, nonelderly households, nondisabled households, non-single-mother households, nonveteran households, and the nonpoor. By contrast, among a number of groups, the median household benefits from two programs—these groups are those with less than a high school diploma or only a high school diploma, the elderly, single-mother households, veteran households, and the poor. The median disabled household receives benefits from three programs. This suggests that program receipt is fairly concentrated. Given the overlapping eligibility criteria for many programs, this degree of concentration is not surprising. The pattern for average benefit receipt among the various demographic groups is quite similar to that for median benefit receipt. Our measure of the average number of cash transfer programs excludes those programs providing in-kind benefits. There is a notable gap between the average number of transfer programs and the average number of cash transfer programs used by single mothers and the elderly.

Next, we investigate what percentage of household income is received from the transfer programs listed in table 2. We do this in two steps. First, in table 4, we show the average benefit received from the different programs by demographic group. These are average monthly benefit amounts among all households in the group. In the final column of table 4, we sum cash transfer income across all the different programs. The elderly receive the largest transfers on average per month ($1,420), followed by the disabled ($1,059) and veterans ($999). Second, in table 5 (p. 124), we tabulate the percentage of total household income that is received from the different transfer programs. In this case, household income is defined as the sum of cash transfers, the value of Food Stamp Program (Supplemental Nutrition Assistance Program) benefits, and other income. While the average U.S. household receives 11 percent of its income from transfer programs (table 5, final row), some groups of households receive (on average) nearly half of their income from these programs. In both tables 4 and 5, we are not imputing values to in-kind assistance, such as Medicare and housing assistance, so these numbers underestimate true total transfer benefits.

We have looked at program participation, benefit levels, and income ratios. By examining transfer programs and their participants in this way, we find that there are certain demographic groups that are particularly dependent on transfer income. We choose to further investigate those groups whose average household (based on the data in table 3) receives benefits from two or more transfer programs (namely, those with less than a high school diploma, the elderly, the disabled, single mothers, and the poor) and also those groups whose ratio of average monthly transfer income to average total monthly income (based on the data in table 5) exceeds 25 percent (namely, those with less than a high school diploma, the elderly, the disabled, and the poor). By using

TABLE 2

Percentage of households receiving government transfer program benefits, by demographic group

	Percent of sample	National School Lunch and Breakfast Programs[a]	Civil Service Retirement System	Earned income tax credit[b]	Food Stamp Program	Hospital and medical care for veterans	Housing assistance	Tenant-based rental assistance/housing choice vouchers (Section 8)
	(percent)							
Less than high school diploma	7.65	24.90	0.74	9.15	22.72	0.63	3.45	2.12
High school graduate, no college	22.66	14.18	1.73	8.40	11.42	1.13	2.16	1.39
Some college or more	69.69	6.74	1.96	5.99	4.07	1.10	0.90	0.61
Elderly	20.64	1.47	6.45	0.62	5.18	1.67	1.35	0.87
Nonelderly	79.36	11.99	0.61	8.52	7.68	0.91	1.39	0.91
Disabled	6.20	16.59	0.55	5.64	32.00	1.86	6.68	5.00
Nondisabled	93.80	9.37	1.90	6.91	5.52	1.02	1.03	0.63
Single mother	7.23	49.28	0.10	30.49	31.10	0.28	6.47	5.22
Other parent	25.70	19.51	0.15	12.81	5.37	0.88	0.46	0.29
Nonparent	46.44	2.02	0.94	2.69	5.31	1.03	1.12	0.58
Veteran	18.32	3.05	5.06	3.49	2.39	4.77	0.55	0.38
Nonveteran	81.68	11.33	1.08	7.51	8.24	0.24	1.57	1.02
Poor	11.93	28.58	0.11	12.06	34.08	0.91	5.85	4.23
Nonpoor	88.07	7.27	2.04	6.03	3.52	1.09	0.78	0.45
All groups	100.00	9.82	1.81	6.82	7.16	1.07	1.38	0.90

TABLE 2 (CONTINUED)

Percentage of households receiving government transfer program benefits, by demographic group

	Medicaid	Medicare	Military retirement	Social Security: Disability Insurance	Social Security: Old-Age and Survivors Insurance	Supplemental Security Income	Temporary Assistance for Needy Families	Unemployment assistance	Veterans service-connected compensation
	(- - - -	- - - -	- - - -	- - - -	- - - - percent - - - -	- - - -	- - - -	- - - -	- - - -)
Less than high school diploma	47.34	41.71	0.18	10.35	36.21	16.08	4.01	2.37	1.20
High school graduate, no college	27.35	34.86	1.09	7.42	31.48	7.66	2.38	2.27	1.43
Some college or more	13.06	19.40	1.72	4.14	17.97	2.78	0.90	1.77	2.16
Elderly	14.86	95.63	3.24	5.85	91.14	6.49	0.30	0.46	3.16
Nonelderly	19.97	6.14	1.00	5.23	4.56	4.49	1.78	2.31	1.59
Disabled	56.10	49.84	1.24	48.69	15.62	33.22	6.31	1.33	3.77
Nondisabled	16.46	22.94	1.48	2.49	22.88	3.03	1.16	1.97	1.79
Single mother	54.47	3.49	0.04	3.20	2.04	8.56	9.11	1.86	0.36
Other parent	23.78	2.66	0.75	2.16	1.65	2.35	1.43	2.68	1.27
Nonparent	12.50	8.48	1.29	7.24	6.56	5.04	0.84	2.18	1.97
Veteran	9.53	43.34	6.85	6.34	42.00	2.28	0.31	1.37	9.47
Nonveteran	21.03	20.41	0.25	5.14	18.04	5.49	1.74	2.06	0.22
Poor	50.86	18.72	0.06	7.37	12.71	15.18	7.09	2.66	0.68
Nonpoor	14.59	25.41	1.65	5.08	23.74	3.51	0.72	1.83	2.08
All groups	18.92	24.61	1.46	5.36	22.43	4.90	1.48	1.93	1.92

[a]The *Survey of Income and Program Participation* asks households specifically about their participation in the National School Lunch and Breakfast Programs, but federal budget sources used for table 1 cover the broader category of child nutrition programs and the Special Milk Program.

[b]Earned income tax credit (EITC) data are missing for about 12 percent of the sample.

Notes: For descriptions of the demographic group variables, see box 1 on p. 119. The sample is limited to metropolitan-based households in which the head is at least 18 years old. There is insufficient detail in the *Survey of Income and Program Participation* to measure the participation in or the benefit value from the refundable (additional) child credit and student assistance, which appear in table 1, as well as distinguish between Medicare Parts A and B. The single mother, other parent, and nonparent households sum to less than 100 percent in the first column of data because the elderly are excluded from these demographic groups (see box 1).

Source: Authors' calculations based on data from wave 4 of the 2004 panel of the U.S. Census Bureau, *Survey of Income and Program Participation*, corresponding to the January–April period of 2005.

TABLE 3

Median and average government transfer program participation, by demographic group

	Median number of programs	Average number of programs	Average number of cash transfer programs
Less than high school diploma	2	2.33	1.05
High school graduate, no college	2	1.67	0.77
Some college or more	0	0.94	0.45
Elderly	2	2.40	1.23
Nonelderly	0	0.91	0.39
Disabled	3	2.93	1.53
Nondisabled	0	1.11	0.51
Single mother	2	2.31	0.87
Other parent	0	0.97	0.32
Nonparent	0	0.63	0.35
Veteran	2	1.48	0.81
Nonveteran	0	1.17	0.52
Poor	2	2.11	0.93
Nonpoor	0	1.09	0.52
All groups	1	1.23	0.57

Notes: For descriptions of the demographic group variables, see box 1 on p. 119. For our measure of participation in cash transfer programs, we exclude programs that supply in-kind benefits to recipients, such as Medicare and housing assistance. We include the Food Stamp Program in our measure of cash transfer program participation.

Source: Authors' calculations based on data from wave 4 of the 2004 panel of the U.S. Census Bureau, *Survey of Income and Program Participation*, corresponding to the January–April period of 2005.

these two different criteria, we end up focusing on the same groups, with the exception of single mothers, who receive 13 percent of their monthly income from transfer programs but are covered by 2.3 programs on average. This discrepancy arises because many single-mother households receive benefits from the child nutrition programs and Medicaid, which are in-kind programs and thus not included in our transfer income calculations.[8]

Group expenditure patterns and inflation rates

We next look at the expenditure patterns of households (or "consumer units") in these five groups: those with less than a high school diploma, the elderly, the disabled, single mothers, and the poor. More specifically, we use data from the *Consumer Expenditure Survey* to investigate whether their expenditure patterns conform to those of the general population. We measure these expenditure patterns by using merged data from the Diary and Interview portions of the survey over the period 1980–2009.[9] The unit of analysis in the *Consumer Expenditure Survey* is the consumer unit—a grouping defined as either a single individual who makes independent consumption decisions, a group of related individuals, or a group of individuals who live together and make joint consumption decisions.[10] We use the term "consumer unit" interchangeably with "household" throughout this article. We define a household as having less than a high school diploma if both the head and spouse have not graduated from high school. We define a household as elderly if either the head or spouse is aged 65 or over. We define a household as headed by a single mother if the household contains children younger than 18 and is headed by an unmarried female aged 18–64. We define a household as disabled if the head or spouse was not working during the past 12 months because he or she was "ill, disabled or unable to work," as stated in the *Consumer Expenditure Survey*'s Codebook. Our definitions here are consistent with the definitions presented in box 1 (p. 119) that we used in our analysis of transfer program participation and the sources of transfer income in tables 2–5.

In table 6, panel A, we report 2009 expenditure shares for our groups of interest, their complements, and the overall population. The number 14.2 in the top row of the column labeled "food" means that among the entire population, 14.2 percent of all expenditures is on food items. We refer to these expenditure shares as the market baskets of households. For all expenditure categories except for housing, these market baskets are based on the out-of-pocket expenditures of households. For example, if a hospital visit was paid for by Medicaid, it would not be included in household

TABLE 4

Average monthly income from government transfer programs, by demographic group

	Social Security (Old-Age, Survivors, and Disability Insurance)	Supplemental Security Income (federal and state amounts)	Food Stamp Program	Military retirement	Civil Service Retirement System	Temporary Assistance for Needy Families (and other welfare amounts)	Unemployment assistance (state and supplemental unemployment benefits)	Veterans service-connected compensation	Earned income tax credit[a]	Total cash transfer income
	(- - - - - - - - - -	- - - - - - - - - -	- - - - - - - - - -	- - - - - - - - - -	- - - - - - - dollars	- - - - - - - - - -	- - - - - - - - - -	- - - - - - - - - -	- - - - - - - - - -	- - - - - - - - - -)
Less than high school diploma	455.35	90.35	45.89	1.02	6.56	15.90	22.23	9.02	15.24	672.86
	(611.98)	(258.84)	(111.76)	(28.04)	(83.20)	(92.88)	(216.13)	(160.42)	(54.00)	(708.69)
High school graduate, no college	455.27	47.66	25.48	12.22	25.18	16.68	22.60	8.14	13.09	632.78
	(732.86)	(202.83)	(92.61)	(132.26)	(230.00)	(325.50)	(187.05)	(100.04)	(47.53)	(899.31)
Some college or more	289.34	19.58	8.80	28.51	40.93	4.80	17.91	12.89	9.09	432.53
	(643.91)	(160.29)	(53.08)	(250.17)	(347.98)	(122.36)	(178.24)	(140.12)	(40.30)	(863.42)
Elderly	1,233.30	32.37	5.56	47.97	116.48	1.51	4.58	20.36	0.74	1,419.88
	(780.44)	(159.48)	(34.01)	(314.48)	(544.01)	(29.75)	(75.29)	(193.29)	(10.59)	(954.03)
Nonelderly	107.23	31.09	17.98	16.15	13.47	10.11	23.13	9.22	13.24	248.08
	(375.16)	(186.23)	(77.14)	(184.95)	(207.59)	(209.80)	(202.11)	(113.32)	(48.37)	(641.03)
Disabled	637.41	228.13	54.62	17.44	11.51	39.73	11.34	42.51	7.89	1,058.50
	(831.69)	(479.31)	(121.48)	(174.55)	(244.42)	(487.77)	(118.46)	(306.90)	(38.15)	(1,100.43)
Nondisabled	319.95	18.35	12.83	23.07	36.27	6.26	19.83	9.47	10.74	460.97
	(650.05)	(130.50)	(65.05)	(221.05)	(315.34)	(147.16)	(186.91)	(113.21)	(43.78)	(833.35)
Single mother	88.75	63.44	88.22	0.61	0.94	46.77	19.76	3.93	49.36	361.78
	(329.08)	(267.67)	(155.40)	(32.04)	(29.57)	(457.54)	(166.08)	(86.11)	(84.32)	(732.37)
Other parent	51.62	18.98	16.29	11.89	3.22	7.54	27.08	6.60	20.48	175.15
	(278.61)	(173.82)	(80.69)	(153.39)	(106.85)	(70.06)	(229.13)	(99.37)	(59.33)	(535.64)
Nonparent	140.88	32.76	7.99	20.93	21.10	5.83	21.47	11.49	3.55	263.42
	(421.72)	(176.53)	(44.04)	(212.64)	(258.95)	(199.26)	(190.94)	(123.76)	(24.70)	(666.16)
Veteran	675.05	16.74	4.30	113.56	111.65	1.94	14.49	55.32	5.16	998.80
	(910.42)	(184.27)	(36.93)	(484.52)	(571.63)	(34.82)	(160.84)	(284.90)	(31.09)	(1,237.90)
Nonveteran	264.40	34.63	17.91	2.34	17.48	9.77	20.38	1.69	11.66	397.04
	(572.84)	(180.14)	(75.92)	(59.28)	(209.44)	(206.69)	(188.10)	(56.67)	(45.49)	(725.66)
Poor	120.97	74.81	82.97	0.35	0.23	25.27	18.22	2.39	19.76	350.06
	(269.19)	(209.13)	(152.48)	(18.04)	(8.11)	(110.17)	(128.83)	(38.54)	(59.93)	(418.38)
Nonpoor	369.25	25.47	6.27	25.75	39.41	6.05	19.45	12.75	9.16	522.60
	(698.72)	(176.06)	(42.56)	(232.52)	(331.60)	(195.44)	(189.61)	(141.90)	(40.19)	(912.58)
All groups	339.63	31.35	15.42	22.72	34.73	8.34	19.30	11.52	10.55	499.98
	(667.15)	(181.03)	(70.61)	(218.45)	(311.46)	(187.42)	(183.42)	(133.87)	(43.44)	(866.00)

[a]Earned income tax credit (EITC) data are missing for about 12 percent of the sample.
Notes: Standard deviations are in parentheses. For descriptions of the demographic group variables, see box 1 on p. 119. The sample is limited to metropolitan-based households in which the head is at least 18 years old. Total cash transfer income is based on about 88 percent of the households for which EITC data are available. Total cash transfer income includes the value of food stamp benefits. Values to in-kind assistance, such as Medicare and housing assistance, are not imputed. There is insufficient detail in the *Survey of Income and Program Participation* to measure the participation in or the benefit value from the refundable (additional) child credit and student assistance, which appear in table 1.
Source: Authors' calculations based on data from wave 4 of the 2004 panel of the U.S. Census Bureau, *Survey of Income and Program Participation*, corresponding to the January–April period of 2005.

TABLE 5

Transfer income as a share of total income, by demographic group

	Average total monthly transfer income (cash transfers and food stamp benefits)	Average total monthly income (including cash transfers and food stamp benefits)	Transfer income as a share of total income
	(- - - - - - - - - *dollars*	- - - - - - - - -)	(*percent*)
Less than high school diploma	672.86	2,150.75	31
	(708.69)	(1,846.15)	
High school graduate, no college	632.78	3,070.68	21
	(899.31)	(2,806.11)	
Some college or more	432.53	5,435.51	8
	(863.42)	(5,641.29)	
Elderly	1,419.88	3,061.06	46
	(954.03)	(3,340.13)	
Nonelderly	248.08	5,023.68	5
	(641.03)	(5,325.44)	
Disabled	1,058.50	2,384.28	44
	(1,100.43)	(2,587.10)	
Nondisabled	460.97	4,756.69	10
	(833.35)	(5,122.98)	
Single mother	361.78	2,785.72	13
	(732.37)	(2,496.91)	
Other parent	175.15	6,363.32	3
	(535.64)	(6,258.03)	
Nonparent	263.42	4,767.55	6
	(666.16)	(5,021.33)	
Veteran	998.80	5,118.47	20
	(1,237.90)	(5,054.67)	
Nonveteran	397.04	4,495.14	9
	(725.66)	(5,019.67)	
Poor	350.06	737.42	47
	(418.38)	(581.62)	
Nonpoor	522.60	5,184.78	10
	(912.58)	(5,146.49)	
All groups	499.98	4,601.77	11
	(866.00)	(5,031.05)	

Notes: Standard deviations are in parentheses. For descriptions of the demographic group variables, see box 1 on p. 119. The calculations in this table are based on the results in table 4.
Source: Authors' calculations based on data from wave 4 of the 2004 panel of the U.S. Census Bureau, *Survey of Income and Program Participation*, corresponding to the January–April period of 2005.

expenditures, but if it was paid for directly by the household, it would be. Expenditure for owner-occupied housing is set equal to the estimated rental value of the property—in keeping with the methodology used by the BLS in the creation of the Consumer Price Index. Entries in panel A of table 6 are in bold if expenditure shares in a given category for a group differ from expenditure shares of the overall population by more than 1 percentage point.

We want to highlight differences in spending in three areas—food, transportation, and health. For food expenditure (table 6, panel A, first column), those with less than a high school diploma, the disabled, single mothers, and the poor all concentrate a higher percentage of expenditures on food than the average consumer. This is in keeping with other research that finds that lower-income households spend a higher portion of their expenditures on food and other necessities. For transportation (fifth column), we see lower expenditure than on average by those with less than a high school diploma, the elderly, the disabled, and the poor. These groups are less likely to have commuting expenses. We find that both the elderly and the disabled spent more on health than the average consumer (sixth column). This pattern is consistent with the weakened health status of these two demographic groups. The nonelderly, single mothers, and the poor spend less on health than the average consumer.

In table 6, panel B, we display annual 2009 expenditure levels by demographic group. We note that total expenditures, as shown in the final column, are higher for those groups that we would expect to have higher income

TABLE 6

Expenditure shares and annual expenditure levels, by demographic group, 2009

	Food	Alcohol	Housing	Apparel	Transportation	Health	Entertainment	Personal care products	Personal care services	Reading materials	Education	Tobacco	Miscellaneous	Total
	(percent)													
A. Expenditure shares														
All groups	14.2	1.0	47.0	3.6	16.3	6.9	5.6	0.7	0.7	0.2	2.4	0.8	0.7	100.0
Less than high school diploma	**16.9**	0.7	**49.0**	**5.1**	**14.5**	6.0	**3.7**	0.8	0.5	0.1	**0.6**	1.6	0.6	100.0
High school graduate or more	14.0	1.0	46.9	3.5	16.4	6.9	5.7	0.7	0.7	0.3	2.5	0.7	0.7	100.0
Elderly	**12.8**	0.8	**50.7**	**2.6**	**12.8**	**12.1**	4.9	0.7	0.7	0.4	**0.4**	0.5	0.7	100.0
Nonelderly	14.6	1.1	46.1	3.8	17.1	**5.6**	5.8	0.7	0.7	0.2	2.8	0.9	0.6	100.0
Disabled	**15.5**	0.9	47.6	3.0	**13.3**	**9.2**	5.3	0.7	0.5	0.2	**0.6**	**2.1**	1.2	100.0
Nondisabled	14.2	1.0	47.0	3.6	16.4	6.7	5.6	0.7	0.7	0.2	2.5	0.7	0.6	100.0
Single mother	**16.2**	0.6	46.4	**5.7**	16.1	**3.9**	5.3	1.1	0.8	0.2	1.7	1.4	0.8	100.0
Non-single-mother	14.2	1.0	47.0	3.5	16.3	7.0	5.6	0.7	0.7	0.3	2.4	0.8	0.6	100.0
Poor	**17.7**	0.6	**48.8**	**4.6**	**12.1**	**5.3**	**4.5**	0.9	0.5	0.2	3.1	1.3	0.6	100.0
Nonpoor	13.8	1.0	46.3	3.6	16.9	7.1	5.8	0.7	0.7	0.3	2.3	0.8	0.6	100.0
B. Annual expenditure levels														
	(dollars)													
All groups	5,953.04	418.46	15,945.73	1,497.62	6,799.08	2,864.70	2,344.95	296.82	280.20	103.36	997.30	331.52	272.49	38,105.27
Less than high school diploma	4,219.20	181.01	10,109.88	1,265.06	3,616.19	1,488.12	920.51	200.47	134.30	29.50	139.03	392.84	138.17	22,834.29
High school graduate or more	6,129.29	442.73	16,627.81	1,527.19	7,175.54	3,026.18	2,508.43	306.57	297.42	112.10	1,099.01	324.30	288.32	39,864.87
Elderly	4,918.05	291.25	12,892.66	987.28	4,915.88	4,634.74	1,874.34	263.19	283.33	142.05	148.26	199.95	276.57	31,827.53
Nonelderly	6,210.37	450.25	16,734.21	1,623.92	7,284.50	2,409.67	2,466.00	304.78	279.41	93.40	1,215.81	365.37	271.50	39,709.19
Disabled	4,490.78	251.34	10,982.11	881.97	3,873.81	2,685.55	1,528.30	199.74	143.13	60.40	175.63	597.32	334.45	26,204.51
Nondisabled	6,038.42	427.79	16,273.70	1,535.58	6,994.20	2,876.99	2,399.50	302.36	289.29	106.22	1,052.22	313.81	268.87	38,878.95
Single mother	5,151.09	176.77	13,589.02	1,801.10	5,132.33	1,248.73	1,691.69	338.73	242.95	54.94	533.27	447.16	248.45	30,656.22
Non-single-mother	6,002.77	433.59	16,079.56	1,480.77	6,892.67	2,955.63	2,381.88	294.02	282.29	106.09	1,023.76	325.07	273.76	38,531.86
Poor	4,030.79	139.37	9,365.43	1,052.20	2,754.12	1,203.07	1,024.09	197.34	112.59	37.03	701.43	295.78	134.76	21,048.01
Nonpoor	6,279.70	475.71	17,140.13	1,616.72	7,683.72	3,214.68	2,630.88	324.46	316.15	120.65	1,057.52	354.17	287.33	41,501.81

Notes: For descriptions of the demographic group variables, see box 1 on p. 119. Expenditure for owner-occupied housing is set equal to the estimated rental value of the property—in keeping with the methodology used by the U.S. Bureau of Labor Statistics in the creation of the Consumer Price Index. Each row's values may not sum to the value in the final column (labeled "total") because of rounding. In panel A, the expenditure shares of specific groups are in bold if they differ from the expenditure shares of the overall population by more than one percentage point.

Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

on average. In particular, total expenditures are higher for those with a high school diploma than those without one, for the nonelderly than the elderly, for the nondisabled than the disabled, for non-single-mothers than single mothers, and for the nonpoor than the poor. In short, individuals in our groups of interest spend less on average than individuals in the remainder of the population.

We measure the inflation of a group as the weighted average of the price changes of the items purchased by households in that group, with the weights depending on the market basket of the group in question. For example, because the elderly spend more on health care than the nonelderly, the price changes in health items get a larger weight in the calculation of the inflation of the elderly than the nonelderly. Given the differences in consumption patterns shown in table 6, panel A (p. 125), we would expect to find differences in inflation experiences if price changes across categories differ dramatically. For example, in a period of rapidly increasing oil prices, we would expect that the inflation experienced by households that commute less, such as the elderly and the disabled, would be lower than that experienced by commuting households, all else being equal.

Before calculating inflation experiences of different groups, we would like to develop some intuition for the results by displaying price changes of goods in different categories. In table 7, panels A and B, we show how prices have changed in the broad expenditure categories displayed in table 6, panels A and B. We show price changes for six different periods: 1980–2010, 1980–90, 1990–2000, 2000–10, 1980–2008, and 2008–10. As noted in the introduction and summary, we divide the period 1980–2010 into 1980–2008 and 2008–10 because of the unusual patterns of price changes and COLAs during the recent recession and subsequent recovery. All price changes are based on August-to-August inflation rates. Panel A of table 7 shows the total price change over the periods, while panel B shows average annual rates during the periods. For example, the 249 percent for food inflation over the period 1980–2010 in panel A of table 7 means that nominal food prices in August 2010 were 249 percent of their August 1980 level. In addition, the average annual rate of food inflation over the period 1980–2010 was 3 percent, as shown in panel B of table 7.

We see in both panels of table 7 that inflation rates have differed across the expenditure categories. For some categories, in particular health, education, and tobacco, price growth has been above the total

TABLE 7

Inflation, by expenditure category, 1980–2010

	Food	Alcohol	Housing	Apparel	Transportation	Health	Entertainment	Personal care products	Personal care services	Reading materials	Education	Tobacco	Miscellaneous	Total
	(percent)													
A. Cumulative price changes														
1980–2010	249	255	265	128	230	514	216	200	273	282	786	1,133	505	262
1980–90	151	149	159	134	143	218	157	159	158	175	249	256	227	158
1990–2000	127	135	131	103	127	159	126	120	135	138	177	212	159	131
2000–10	130	127	127	93	126	148	109	105	128	117	178	209	140	126
1980–2008	245	246	268	128	246	482	217	197	266	273	715	823	484	263
2008–10	101	104	99	100	94	107	100	101	103	103	110	138	104	100
B. Annual average inflation														
1980–2010	3	3	3	1	3	6	3	2	3	4	7	8	6	3
1980–90	4	4	5	3	4	8	5	5	5	6	10	10	9	5
1990–2000	2	3	3	0	2	5	2	2	3	3	6	8	5	3
2000–10	3	2	2	–1	2	4	1	0	3	2	6	8	3	2
1980–2008	3	3	4	1	3	6	3	2	4	4	7	8	6	4
2008–10	1	2	0	0	–3	3	0	1	1	2	5	17	2	0

Note: All price changes are based on August-to-August inflation rates.
Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, Consumer Price Index Database.

price growth (as shown in the final column) during all three decades of the 1980–2010 period. In contrast, apparel inflation has been lower than the overall price growth during all three decades. Transportation price growth has been lower than or equal to total price growth in all of the periods we consider. Because of these patterns, we would expect groups that concentrate high portions of consumption on health, education, and tobacco to have experienced higher inflation than the average consumer, while groups that concentrate high portions of spending on apparel and transportation would have experienced lower inflation.

Now, we combine the expenditure share data and the price change data to calculate group inflation. We calculate group inflation in two ways. Our first inflation calculation is based on the annual market basket consumed by a particular group. The inflation rate for a group in a particular month is calculated as the year-over-year price change of the market basket consumed by that group in the prior year. For example, inflation for the elderly in August 2010 is equal to the price change between August 2009 and August 2010 of the market basket purchased by the elderly in 2009.[11] Put differently, inflation is the weighted average price change of the goods and services purchased by the elderly, with the weights being the elderly's expenditure shares (as displayed in table 6, panel A, p. 125). We label such calculations "annual-weighted inflation." This differs from the way in which the official CPI is calculated because the official CPI uses weights that are fixed over a period longer than a year and are derived from expenditures across multiple years. For example, the CPI from January 2006 through December 2007 is based on the 2003 and 2004 market basket. Our second inflation calculation follows the BLS's methodology as closely as we are able (U.S. Bureau of Labor Statistics, 2007).[12] For this second measure, we only tabulate inflation from 1987 onward because earlier inflation data would require the use of older *Consumer Expenditure Survey* data (in particular that for 1972–73) than we have used. We label such calculations "fixed-weighted inflation."[13]

In table 8, panel A, we show annual-weighted inflation calculations, and in panel B, we show fixed-weighted inflation calculations. We show cumulative inflation experiences based on inflation during the month of August. We choose August because many of the COLAs are based on year-over-year third-quarter inflation. In table 8, panel A (first row and first column of data), we show that for the overall population, prices were 255 percent of their August 1980 level in August 2010. This does *not* mean that a fixed set of goods that cost $100 in 1980 costs $255 in 2010 because our calculations of inflation are based on a market basket that is redetermined every year.

Over the 1980–2010 period, the highest levels of inflation have been experienced by the elderly, followed by the disabled, the poor, and those with less than a high school diploma, as shown in table 8, panel A (as well as in panel B over the 1987–2010 period). This pattern is due in part to the findings presented in panel A of table 6 (p. 125) that the elderly and the disabled spend more than on average in the health category, which had quickly growing prices, while those with less than a high diploma and the poor spend less than on average in the transportation category, which had slowly growing prices. This general pattern persists, more or less, across the different periods displayed in table 8, panel A. This finding is consistent with other research that has focused on the elderly as a group that has faced high inflation (Hobijn and Lagakos, 2005; and Amble and Stewart, 1994).

Based on the calculations using annual weights in panel A of table 8, we note that over the 30-year period from 1980 through 2010, inflation faced by the elderly has been 15 percentage points higher than that experienced by the overall population. Inflation faced by the elderly has been higher in each of the three decades displayed in panel A of table 8 as well. We generally find smaller gaps between the inflation of the poor, those with less than a high diploma, and the disabled and that of the overall population. We also find that single mothers have experienced slightly lower inflation than the overall population. The results using fixed weights, in panel B of table 8, are similar. Note that the numbers in the first row of panel B of table 8 are smaller than the numbers in the first row of panel A of table 8 because we are measuring cumulative inflation over a shorter period in panel B.

In the final column of both panels A and B of table 8, we show cumulative August-to-August inflation according to the official CPI-U. Our measure of inflation for "all" over the period 1987–2010 in panel B of table 8 (190 percent in the first row and first column of data) should be close to the official CPI-U over the same period (191 percent in the first row and final column) because for this data point we are using the same BLS data and methodology. We would expect our inflation measure for "all" over the period 1980–2010 in panel A of table 8 (255 percent in the first row and first column) to be lower than the official CPI-U over the same period (262 percent in the first row and final column) because we are updating market baskets more quickly than the CPI-U and taking into account the fact that households may change their behavior in response to rising prices by purchasing more of those goods and services whose prices are increasing less quickly.

Group inflation and program COLAs

Our next goal is to compare the inflation experiences of different groups to increases in benefit payments. Benefit payment increases arise either because a program has an explicit cost-of-living adjustment or because legislators enact increased benefit amounts. We focus on four programs—Social Security, SSI, TANF, and SNAP.

Social Security and SSI

We begin by looking at the Social Security and SSI COLA and the inflation experiences of the elderly and the disabled. Social Security and SSI benefits (for both the elderly and the disabled) have been indexed to the (seasonally unadjusted) CPI-W since 1975. Benefits for the Civil Service Retirement System, military retirement, and veterans service-connected compensation are all indexed in the same way.

In table 9, panel A, we show the increase in the CPI-W in the first column of data. The number 256 in the first row and first column of data means that, according to the CPI-W, prices in August 2010 were 256 percent of their nominal August 1980 value. In the next two columns, the numbers displayed for the various periods are the same as those measuring the inflation faced by the elderly and the disabled with annual weights and fixed weights in table 8, panels A and B, respectively.

From table 9, panel A, we see that for both annual-weighted and fixed-weighted inflation measures, the inflation experienced by the elderly has been almost always higher than the CPI-W, both over the entire period and for each of the three decades covered by the data. Over the entire 30-year period, based on annual weights, elderly inflation has been 14 percentage points above the CPI-W. For each of the three decades presented in the part of panel A of table 9 using annual weights, the gap has been between 2 percentage points and 5 percentage points. Because individuals tend to benefit from the program for a number of years (the life expectancy of an American 65 years old in 1980 was 16.4 additional years),[14] these decade-long gaps lead to declines in the purchasing power for the same individual. For the disabled, the inflation experienced by the group has also tended to be higher than aggregate inflation for both the

TABLE 8

Cumulative inflation experiences, by demographic group

	All	Less than high school diploma	High school graduate or more	Elderly	Nonelderly	Disabled	Nondisabled	Single mother	Non-single-mother	Poor	Nonpoor	Official CPI-U
	(- - - -	- - - -	- - - -	- - - -	- - - -	- - percent	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -)
A. Annual weights												
1980–2010	255	261	255	270	253	262	255	247	256	262	254	262
1980–90	156	157	156	160	156	158	156	155	156	158	156	158
1990–2000	129	130	129	132	129	129	129	127	130	129	129	131
2000–10	126	128	126	128	126	128	126	125	126	129	126	126
1980–2008	255	262	255	269	253	260	255	247	256	261	254	263
2008–10	100	100	100	100	100	100	100	100	100	100	100	100
B. Fixed weights												
1987–2010	190	194	190	198	188	197	190	187	190	195	189	191
1987–90	114	115	114	115	114	115	114	114	114	115	114	115
1990–2000	131	131	131	133	131	133	131	130	131	132	131	131
2000–10	127	129	127	129	126	129	127	126	127	129	126	126
1987–2008	190	195	190	198	189	197	190	187	190	195	189	191
2008–10	100	100	100	100	100	100	100	100	100	100	100	100

Notes: For descriptions of the demographic group variables, see box 1 on p. 119. CPI-U means Consumer Price Index for All Urban Consumers. The cumulative inflation experiences are based on inflation during the month of August. Please see the text for further details on inflation based on annual and fixed weights. The different weighting methodologies do not apply to the calculations for the official CPI-U, which is created by the U.S. Bureau of Labor Statistics using one weighting methodology.

Sources: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, Consumer Price Index Database and *Consumer Expenditure Survey*.

TABLE 9

Consumer Price Index, benefit adjustments, and inflation experiences of select demographic groups

A. Social Security and SSI and the elderly and the disabled

	Official CPI-W	Elderly	Disabled	Social Security/ SSI COLA
	(- - - - - - - - - -	percent	- - - - - - - - - -	)
Annual weights				
1980–2010	256	270	262	265
1980–90	155	160	158	152
1990–2000	130	132	129	133
2000–10	126	128	128	131
1980–2008	257	269	260	250
2008–10	100	100	100	106
Fixed weights				
1987–2010	189	198	197	198
1987–90	115	115	115	113
1990–2000	130	133	133	133
2000–10	126	129	129	131
1987–2008	190	198	197	187
2008–10	100	100	100	106

B. AFDC/TANF and single mothers

	AFDC/TANF maximum in Alabama	AFDC/TANF maximum in Connecticut	AFDC/TANF maximum in Illinois	Official CPI-W	Single mother
	(- - - - - - - - - -	- - - - - - - - - -	percent	- - - - - - - - - -	)
Annual weights					
1980–2010	182	143	150	256	247
1980–90	100	137	127	155	155
1990–2000	139	99	103	130	127
2000–10	131	106	114	126	125
1980–2008	182	143	150	257	247
2008–10	100	100	100	100	100
Fixed weights					
1987–2010	182	124	126	189	187
1987–90	100	119	107	115	114
1990–2000	139	99	103	130	130
2000–10	131	106	114	126	126
1987–2008	182	124	126	190	187
2008–10	100	100	100	100	100

C. SNAP and the disabled, single mothers, poor, and those with less than a high school diploma

	Thrifty Food Plan	CPI-food	Official CPI-W	Disabled	Single mother	Poor	Less than high school diploma	SNAP (food stamp) monthly maximum
	(- - - - - - - - - -	- - - - - - - - - -	- - - - - - - - - -	percent	- - - - - - - - - -	- - - - - - - - - -	- - - - - - - - - -	)
Annual weights								
1980–2010	250	249	256	262	247	262	261	320
1980–90	149	151	155	158	155	158	157	158
1990–2000	128	127	130	129	127	129	130	129
2000–10	131	130	126	128	125	129	128	157
1980–2008	260	245	257	260	247	261	262	259
2008–10	96	101	100	100	100	100	100	123
Fixed weights								
1987–2010	202	193	189	197	187	195	194	246
1987–90	120	117	115	115	114	115	115	122
1990–2000	128	127	130	133	130	132	131	129
2000–10	131	130	126	129	126	129	129	157
1987–2008	210	190	190	197	187	195	195	200
2008–10	96	101	100	100	100	100	100	123

Notes: For descriptions of the demographic group variables, see box 1 on p. 119. CPI-W means Consumer Price Index for Urban Wage Earners and Clerical Workers; CPI-food means the Consumer Price Index for all food. COLA means cost-of-living adjustment. SSI means Supplemental Security Income. AFDC means Aid to Families with Dependent Children, and TANF means Temporary Assistance for Needy Families. SNAP means Supplemental Nutrition Assistance Program (Food Stamp Program). The Thrifty Food Plan is the basis for food stamp allotments. Please see the text for further details on inflation based on annual and fixed weights. The different weighting methodologies only apply to the inflation calculations for the demographic groups.

Sources: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, Consumer Price Index Database and *Consumer Expenditure Survey*; and data on benefit determination from the U.S. Social Security Administration (panel A), U.S. Department of Health and Human Services (panel B), and U.S. Department of Agriculture (panel C).

annual-weighted and fixed-weighted measures, although to a lesser degree (for example, by 6 percentage points over the entire period for the annual-weighted measure). Like the elderly on Social Security Old-Age and Survivors Insurance, disabled individuals tend to benefit from the Social Security Disability Insurance and SSI programs for long durations—with the average stay on disability lasting over ten years (Rupp and Scott, 1995).

The comparison of inflation experiences of the elderly and the disabled with the actual COLAs implemented by the Social Security and SSI programs is more complicated. Over the period 1976–83, the Social Security/SSI COLAs were based on increases in the CPI-W from the first quarter of the prior year to the first quarter of the current year and became effective with June benefits paid to recipients in July. After 1983, the COLAs were based on increases in the CPI-W from the third quarter of the prior year to the third quarter of the current year and became effective with December benefits paid in January. Figure 1 shows August-over-August growth in the CPI-W, Social Security/SSI COLA, and inflation faced by the elderly. There are three notable features of the Social Security/SSI COLA relative to the CPI-W. First, CPI-W increases are reflected in the COLA with a lag because the COLA has been implemented one quarter after the price change is measured. Second, there was no COLA in 1983; in other words, benefits in August 1983 were the same as benefits in August 1982. This is due to the shift, beginning in 1983, from implementing COLAs in July to implementing COLAs in the following January (that is, the 1983 COLA was implemented in January 1984). Third, recent Social Security/SSI COLAs have been somewhat unusual. The COLA for 2008—first paid in January 2009—was 5.8 percent. Prices in 2008:Q3 were 5.8 percent above their 2007:Q3 level. The magnitude of this increase was in part due to the timing of the COLA determination. Energy prices spiked over the summer of 2008. For all of 2008, CPI-W inflation was 4.1 percent, but the COLA was based on the 2008:Q3 measurements. The COLA for 2009—first paid in January 2010—was zero because prices fell between 2008:Q3 and 2009:Q3 and the COLA cannot be negative. This fall was due in part to the temporary nature of the energy price spike. The COLA for 2010, payable in January 2011 (not shown in figure 1), was also zero because, although prices increased modestly (1.5 percent) between 2009:Q3 and 2010:Q3, they were still about half a percentage point below their 2008:Q3 level. In effect, recipients were compensated beginning in January 2009 for the inflation experienced between 2009:Q3 and 2010:Q3. Because of all these factors, the CPI-W and the Social Security/SSI COLA have differed modestly over this period.

We divide our comparison of the Social Security/SSI COLA with elderly inflation into the periods 1980–2008 and 2008–10 because the forces at work in these two eras differ. In the 1980–2008 period, elderly inflation was above the Social Security/SSI COLA by 19 percentage points (see table 9, panel A, fifth row, p. 129). This is in part due to the following factors: the gap between elderly inflation and overall inflation, the lack of a COLA in 1983, and the fact that the price increases in 2008 had not yet been incorporated into the COLA. In 2008–10, the Social Security/SSI COLA was higher than the inflation faced by the elderly. This is due to the large COLA in January 2009 and the fact that the January 2010 COLA could not be negative. The pattern for the disabled is similar, although the gap in the 1980–2008 period is smaller.

Overall, the inflation experienced by the elderly and the disabled has generally been higher than the price index used to adjust their most substantial income support benefits. However, the path for the actual Social Security/SSI COLA has differed from that for the index it tracks because of idiosyncrasies in the determination of the Social Security/SSI COLA.

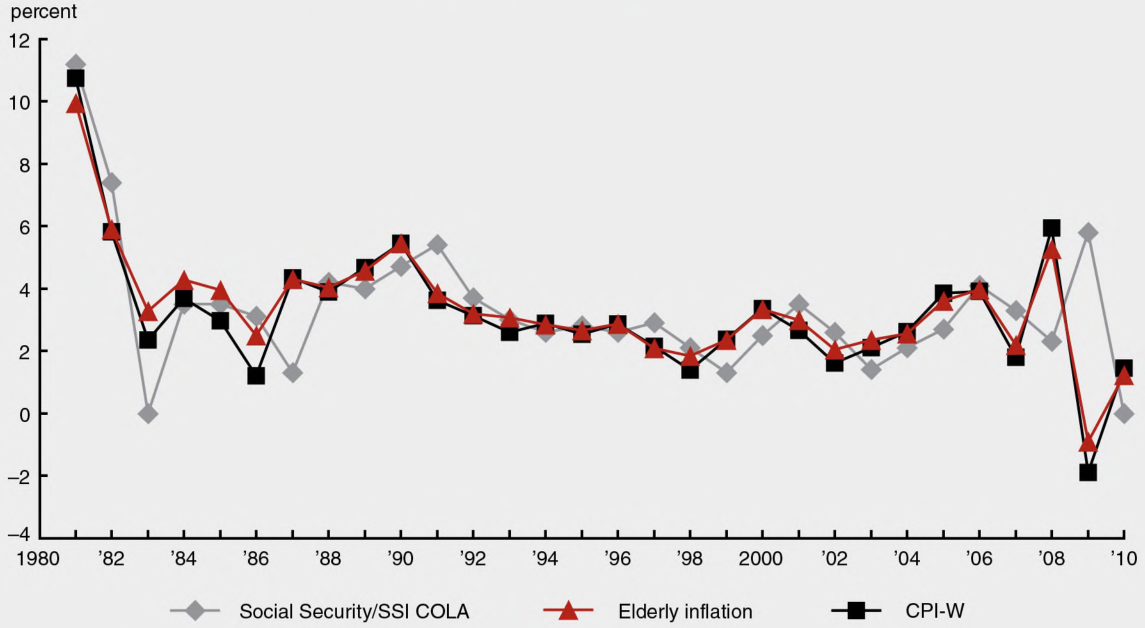
Temporary Assistance for Needy Families

Benefit payments for the Temporary Assistance for Needy Families program (which replaced the Aid to Families with Dependent Children in 1997) are set by the states. States set maximum benefit payments for families of different compositions, and subtract some portion of family income to set the actual benefit payment for a given family. In table 9, panel B (p. 129), we compare nominal increases in maximum monthly AFDC/TANF benefits for a family of three in three states—Alabama, Connecticut, and Illinois—with the inflation experiences of single mothers, based on both annual and fixed weights. We choose these three states because one was a relatively high-benefit state in 1980 (Connecticut's maximum benefit was $475), one was a moderate-benefit state (Illinois's maximum was $288), and one was a low-benefit state (Alabama's maximum was $118). While states determine benefit levels, TANF payments are partially funded by federal block grants that have been fixed in nominal terms since they were established in 1996.

If we compare the increases in AFDC/TANF benefits with the inflation experiences of single mothers based on annual weights, we find that while single mothers were facing prices in 2010 that were 247 percent of their 1980 level, benefits in these three states were between 143 percent and 182 percent of

FIGURE 1

Social Security/SSI COLA, elderly inflation, and CPI-W



Notes: SSI means Supplemental Security Income. COLA means cost-of-living adjustment. CPI-W means Consumer Price Index for Urban Wage Earners and Clerical Workers. August-over-August growth is displayed for all data.
Sources: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, Consumer Price Index Database and *Consumer Expenditure Survey*; and data on benefit determination from the U.S. Social Security Administration.

their 1980 level (see table 9, panel B, first row, p. 129). Growth in the price of the market basket of single mothers was 65 percentage points above the growth in benefits in the state with the largest percentage growth in benefits among the three selected—Alabama. We also see large gaps, particularly for Connecticut and Illinois, when we investigate the 1987–2010 period and use fixed weights. These gaps between benefit growth and price growth are far larger than that between the Social Security/SSI COLA and elderly inflation, and represent substantial erosion in the purchasing power of program beneficiaries. These three states are fairly representative of the 50 states. In no state did the value of benefits keep up with the annual-weighted price increases faced by single mothers over the 1980–2010 period. This erosion in the real value of welfare benefits has been noted elsewhere (for example, Schott and Levinson, 2008). For 1990–2000 and 2000–10, growth in benefit payments in Alabama (the low-benefit state) slightly exceeded the price growth faced by single mothers (see table 9, panel B, third and fourth rows). However, the maximum benefit in Alabama had been unchanged between 1980 and 1990 (see table 9, panel B, second row).

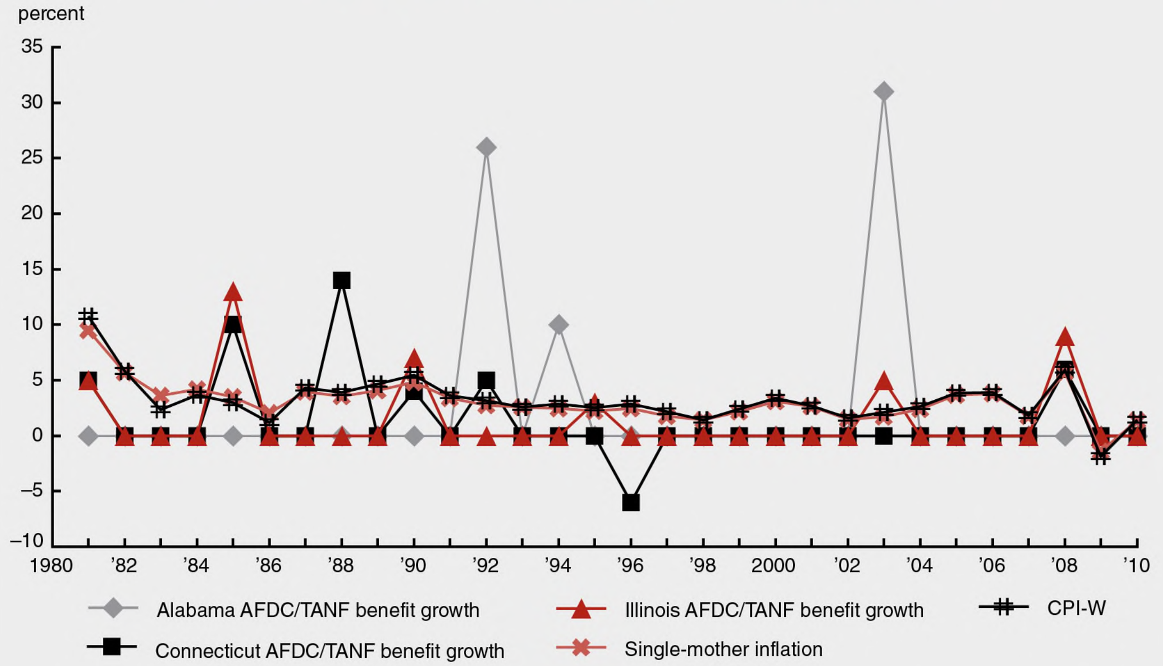
In figure 2, we show August-over-August increases in AFDC/TANF benefits in the three states, overall inflation (as measured by the CPI-W), and single-mother inflation. In most years, benefits have been unchanged, but there have been occasional changes. For the AFDC/TANF population, the duration of benefit receipt differs before and after the implementation of the TANF program in 1997, since federal funding for TANF recipients is limited to 60 months. Prior to welfare reform in 1996, over 50 percent of the caseload was expected to stay on the program for over a decade (Rupp and Scott, 1995). The real erosion in welfare benefits translates into both lower real benefits for individuals who enter AFDC/TANF at later dates and a decline in the purchasing power of benefits for an individual during her stay on AFDC/TANF.

Supplemental Nutrition Assistance Program

In panel C of table 9 (p. 129), we compare increases in monthly maximum benefits from the Supplemental Nutrition Assistance Program (formerly called the Food Stamp Program) with price increases faced by those with less than a high school diploma, the disabled, single mothers, and the poor—all based on both annual and fixed weights. Individuals in all of these groups are heavily represented among SNAP recipients (see the Food Stamp Program column in

FIGURE 2

AFDC/TANF benefit growth in select states, single-mother inflation, and CPI-W



Notes: AFDC means Aid to Families with Dependent Children; TANF means Temporary Assistance for Needy Families, and it replaced the AFDC in 1997. CPI-W means Consumer Price Index for Urban Wage Earners and Clerical Workers. August-over-August growth is displayed for all data.
Sources: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, Consumer Price Index Database and *Consumer Expenditure Survey*; and data on benefit determination from the U.S. Department of Health and Human Services.

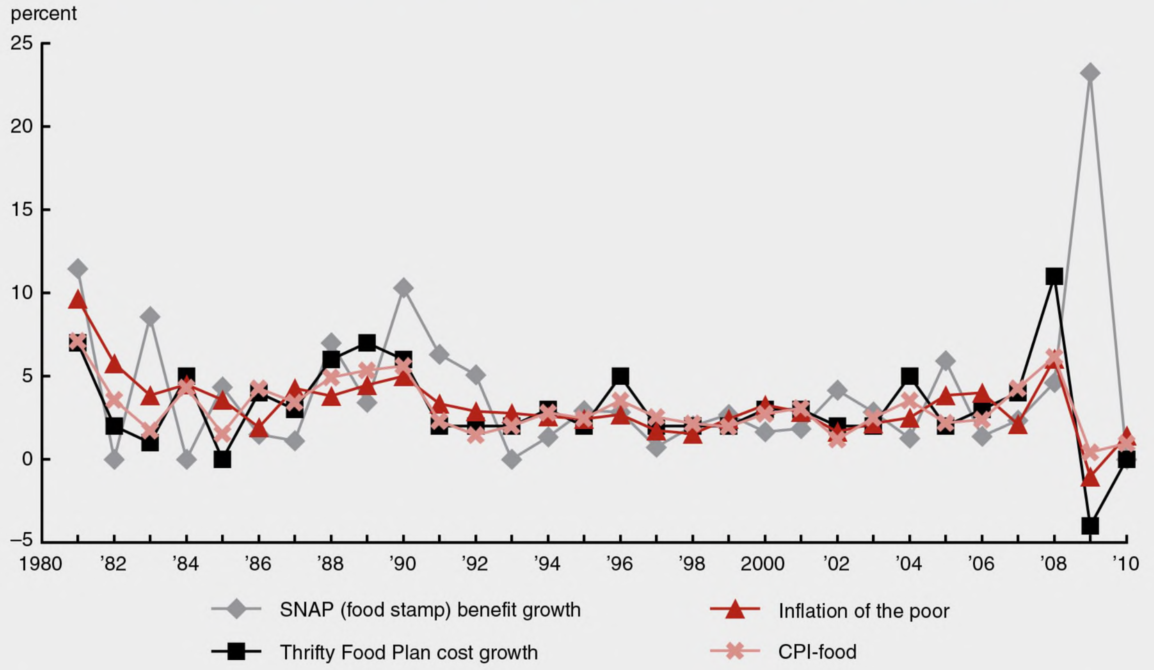
table 2, pp. 120–121). As noted in table 1 (pp. 115–117), SNAP maximum benefits are currently indexed to increases in the cost the U.S. Department of Agriculture's Thrifty Food Plan (TFP). In particular, benefits are based on the cost of a low-cost nutritious diet for a family of four with one child aged 6–8 and one child aged 9–11. June-to-June increases in prices are reflected in October SNAP benefits.

In table 9, panel C (p. 129), we display increases in the cost of the TFP in the first column of data. The fourth through seventh columns of data display price increases for our groups of interest. The increases in the cost of the TFP are slightly above the annual-weighted price increases faced by single mothers over the 1980–2010 period, while the increases in the cost of the TFP are slightly below the inflation experienced by the other groups in panel C. For the 1980–90 period, when food price growth overall (as measured by the Consumer Price Index for food, or CPI-food, and displayed in the second column) was below total inflation (displayed in the third column), all groups experienced inflation that was higher than the growth in the cost of the TFP. This is by design, in that growth in benefits is meant to match increases in the price of food rather than increases in the cost of all items.

We graph annual August-over-August increases in SNAP benefits, the cost of the TFP, the cost of food (as measured by the CPI-food), and the price of the market basket consumed by the poor in figure 3. Two notable patterns emerge from the figure. First, TFP inflation is more volatile than overall food inflation (as measured by the CPI-food). This is due to greater weighting in the TFP toward vegetables, milk products, and fruit, whose prices tend to be less stable than those of other foods and food away from home (McGranahan, 2008; and Carlson et al., 2007). Second, SNAP benefit growth and TFP cost growth have differed quite substantially at times. In fact, over the period 1980–2010, these measures have been *negatively* correlated. This differential is due to the four-month lag in implementing benefit changes and to frequent policy changes in the methods of determining maximum benefits. The current method of indexing benefits was first put in place in October 1996, based on June 1996 prices.[15] Prior to that, inflation adjustments had been a done in a variety of different ways. Since the Food Stamp Act of 1964,

FIGURE 3

SNAP benefit growth, Thrifty Food Plan cost growth, inflation of the poor, and food inflation



Notes: Supplemental Nutrition Assistance Program (SNAP) was previously known as the federal Food Stamp Program (and it is still commonly referred to by this name). The Thrifty Food Plan is the basis for food stamp allotments. CPI-food means Consumer Price Index for all food. August-over-August growth is displayed for all data.

Sources: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, Consumer Price Index Database and *Consumer Expenditure Survey*; and data on benefit determination from the U.S. Department of Agriculture.

food stamps have been indexed annually, indexed semi-annually, and frozen. Benefits have been set from 99 percent to 103 percent of the cost of the Thrifty Food Plan. For example, from October 1992 until October 1996, food stamp benefits were set at 103 percent of the cost of the TFP. As another example, food stamp benefits were fixed between January 1981 and September 1982, and the benefit adjustment for October 1, 1982, was based on 21 months of price changes.

By contrast, in 2010, the food stamp maximum was 123 percent of its 2008 level, although the cost of the TFP was 96 percent of its 2008 level in 2010 (table 9, panel C, sixth row, p. 129). This disparity is due to a provision of the American Recovery and Reinvestment Act of 2009 that set benefits beginning in April 2009 at 113.6 percent of the cost of the TFP as of June 2008. Because of the ARRA increase, SNAP benefit growth exceeded the inflation faced by all population groups over the entire 1980–2010 period—in particular, the 2008–10 period.

SNAP COLAs will also be unusual going forward. Under current legislation, SNAP benefits are set to remain at 113.6 percent of the June 2008 TFP cost until October 2013 unless food inflation is so high that 100 percent of the contemporaneous TFP cost exceeds 113.6 percent of the June 2008 TFP cost prior to that date. In other words, from now until October 2013, unless inflation is very high, there will be no benefit increases. In October 2013, benefits will revert to being set at 100 percent of the June 2012 TFP cost. Assuming total TFP inflation between June 2008 and October 2013 is less than 13.6 percent (about 2.6 percent per year), benefits will fall in October 2013. This schedule for future benefit adjustments has been changed twice since the passage of the ARRA. Originally, the provision was going to end whenever the TFP cost exceeded 113.6 percent of the June 2008 level. It was then set to end in March 2014 and is now set to end in October 2013. These changes in the timing of the added benefits' phaseout are akin to what has been seen in other periods where food stamp benefit adjustments were subject to frequent policy shifts. In general, the relationship of the cost of the TFP and maximum SNAP allotments has been influenced by policy decisions. This relationship was altered through the ARRA and through two additional pieces of legislation since the ARRA's passage, as well as numerous times prior to 2009.

Reviewing the four transfer programs

The relationship between the experience of program recipients and the computation of benefit levels has been quite different for Social Security and SSI, TANF, and SNAP. With the exception of a change in the timing of COLA determination between 1982 and 1983, the Social Security/SSI COLA has been calculated in a consistent manner. As a result, for the most part, the Social Security/SSI COLA has been close to the inflation measure, the CPI-W, it is intended to track. However, the inflation experiences of both the elderly and the disabled have been generally higher than the inflation of the population covered by the CPI-W.

The gap between the Social Security/SSI COLA and the inflation of the elderly and the disabled pales in comparison with the gap between the growth of TANF benefits and the inflation faced by single mothers. TANF beneficiaries have seen substantial declines in the purchasing power of their benefits. Although the inflation faced by single mothers has been moderately below inflation for the CPI-W population, the growth in TANF benefits has been far below the inflation faced by single mothers. This gap is so large because neither the block grants from the federal government to the states nor the state benefits themselves are indexed. As a result, the growth in nominal TANF benefits at the state level has been modest and uneven.

For SNAP benefit recipients, inflation over the entire period has been close to growth in the cost of the TFP in part because food inflation has been close to overall inflation. However, SNAP benefit growth and TFP cost growth have diverged because of policy decisions. SNAP benefit levels and the relationship between these benefit levels and the cost of the TFP have been policy levers that are frequently used. Because of a major increase in benefits enacted as part of the ARRA, benefit increases far exceed the inflation of the groups that depend on SNAP. However, the history of the ARRA benefits is also indicative of the frequency with which SNAP benefits are altered—the timing of the phaseout of the added ARRA benefits has been changed twice since the ARRA passed.

Conclusion

We compare the inflation indexation used in government transfer programs with the inflation experiences of households that are dependent on those programs for income support.

We find that the inflation experienced by different demographic groups differs from aggregate inflation because of differences in consumption patterns across the demographic groups and differences in price changes across expenditure categories. Demographic groups that concentrate a higher portion of their spending in categories whose prices have grown rapidly over the past three decades, such as health care, have experienced higher inflation than demographic groups that concentrate a higher portion of their spending in categories whose prices have grown more slowly, such as transportation. Because of their high demand for health care and low commuting costs, elderly households have experienced the highest inflation of all the groups we investigate.

We also find that the evolution of transfer program benefits has differed substantially across the four programs we investigate. Social Security and SSI benefit growth has been moderately lower than the inflation experiences of the elderly and the disabled because of the consumption patterns of the elderly and the disabled. However, TANF benefit growth has been far below the inflation experienced by single mothers because of the absence a routine COLA for most state-level benefits; and SNAP benefit growth has diverged from the inflation experienced by its beneficiaries because of frequent changes in the way the SNAP COLA has been calculated.

Much of the policy debate concerning COLAs has revolved around the Social Security program. This is in part due to the high inflation experienced by the elderly and in part due to the fact that Social Security is the single largest income support program, representing 31 percent of all federal expenditures on payments to individuals in 2010.[16] Because the elderly have experienced higher inflation than the overall population, their inflation experiences have exceeded inflation as measured by the CPI-W, the index upon which increases in Social Security benefits are based. Given the gap between inflation experienced by the elderly and the Social Security COLA, the elderly have experienced a decline in their ability to purchase their preferred market basket, even in the presence of a fully indexed benefit. Using an alternative COLA that indexed Social Security benefits to the inflation faced by the elderly could eliminate this gap. However, such a policy change would be extremely costly. Researchers at the Federal Reserve Bank of New York (Hobijn and Lagakos, 2003) estimated the potential costs of using a CPI based on the consumption patterns of the elderly to adjust Social Security benefits. According to this research, had an elderly-specific index been adopted in 1984, benefits would have been 3.84 percent higher than they actually were in 2001. The New York Fed researchers anticipated that changing to an elderly-specific index in 2003 would likely have increased future benefit levels and have rendered the Social Security trust fund insolvent five years sooner than projected at the time.

Alternatively, changing the Social Security COLA to one that led to more modest increases in benefits, as proposed by the National Commission on Fiscal Responsibility and Reform, would magnify the gap between the inflation of the elderly and the Social Security COLA.[17] At the same time, such a change would relieve some budget pressures.

Past attempts to change the Social Security COLA to one that reflected the purchasing habits of the elderly have not gotten much traction. Legislation to change the Social Security COLA has been introduced in every Congress since the 105th in 1997–98, but this legislation has never made it to the floor of either chamber. Legislation has also been introduced in the current Congress. The National Commission on Fiscal Responsibility and Reform's proposal to use a chain-weighted Consumer Price Index (in particular, the chain-weighted Consumer Price Index for All Urban Consumers, or C-CPI-U) has not yet been incorporated into any legislative proposal, and such a change was not included in President Obama's 2012 budget proposals. However, this change has been incorporated into some of the broad-based proposals to address the federal deficit.

NOTES

[1]Social Security is officially referred to as the federal Old-Age, Survivors, and Disability Insurance Program, or OASDI.

[2]Chain-weighting is a method of measuring inflation that takes into account the fact that people tend to buy less of things whose prices have increased a lot and instead buy more of substitutes whose prices have risen less. Different inflation measures are discussed in more detail later in this article.

[3]Social Security and SSI are administered by the U.S. Social Security Administration; TANF is administered by the U.S. Department of Health and Human Services; and SNAP is administered by the U.S. Department of Agriculture, Food and Nutrition Service.

[4]These are our calculations based on data from White House, Office of Management and Budget (2011a, b).

[5]These adjustments are automatic except for the case of veterans' benefits. Congress enacts legislation every year that sets the COLA for veterans' benefits equal to that for Social Security benefits. This legislation tends to pass unanimously.

[6]Market baskets refer to evolving selections of goods and services purchased by individuals that are used to track inflation in an economy or specific market.

[7]For alternative discussions on the use of indexation in the federal government, see Congressional Budget Office (1981, 2010).

[8]Our criteria do not lead us to look at the expenditure patterns of veterans. However, even if we had wanted to do so, we would not have been able to. The only question related to veteran status in the *Consumer Expenditure Survey* is one that asks the amount of income from workers' compensation and veterans' benefits combined. Workers' compensation is a larger program paying out $55 billion in annual benefits in 2007 (U.S. Census Bureau, 2010) versus $32 billion in veterans service-connected compensation (see table 1 on pp. 115–117 of this article for more on the latter). We do not include workers' compensation in the list of programs discussed in this article because it is largely funded by private insurance carriers and employers' self-insurance and not by the federal government.

[9]In McGranahan and Paulson (2006), we detail how the data are merged to create these expenditure patterns.

[10]For the official BLS definition of "consumer unit," see www.bls.gov/cex/csxfaqs.htm#q3.

[11]Doing the calculation in this way ensures that the inflation measure is not influenced by seasonal patterns in expenditures or prices.

[12]We cannot perfectly mirror the BLS's calculation because we lack some of the information needed to do so. Area information is missing in the public use data. In addition, some prices are not publicly released.

[13]These weights are not fixed over the entire sample, but they are fixed for longer than a year.

[14]See www.cdc.gov/nchs/data/hus/hus2009tables/Table024.pdf.

[15]The current method of indexing SNAP (food stamp) benefits was enacted in the Personal Responsibility and Work Opportunity Reconciliation Act of 1996. See www.fns.usda.gov/snap/rules/Legislation/timeline.pdf for details on the Food Stamp Program's evolution.

[16]This is our calculation based on data from White House, Office of Management and Budget (2011b).

[17]It would be worthwhile to compare the Social Security COLA with increases in elderly inflation by using the type of chain-weighted measures that the National Commission on Fiscal Responsibility and Reform has proposed. While chain-weighted inflation experienced by the elderly will most likely be lower than the measures of elderly inflation we present, the rapid increases in health care costs will still lead to chain-weighted inflation experienced by the elderly being greater than chain-weighted aggregate inflation.

Amble, Nathan, and Kenneth Stewart, 1994, "Experimental price index for elderly consumers," *Monthly Labor Review*, Vol. 117, No. 5, May, pp. 11–16.

Carlson, Andrea, Mark Lino, WenYen Juan, Kenneth Hanson, and P. Peter Basiotis, 2007, *Thrifty Food Plan, 2006,* U.S. Department of Agriculture, Center for Nutrition Policy and Promotion, report, No. CNPP-19, April.

Congressional Budget Office, 2010, "Using a different measure of inflation for indexing federal programs and the tax code," Economic and Budget Issue Brief, Washington, DC, February 24.

__________, 1981, *Indexing with the Consumer Price Index: Problems and Alternatives*, report, Washington, DC, June.

Hanson, Kenneth, and Margaret Andrews, 2008, "Rising food prices take a bite out of food stamp benefits," Economic Information Bulletin, U.S. Department of Agriculture, Economic Research Service, No. 41, December.

Haveman, Robert, Karen Holden, Kathryn Wilson, and Barbara Wolfe, 2003, "Social security, age of retirement, and economic well-being: Intertemporal and demographic patterns among retired-worker beneficiaries," *Demography*, Vol. 40, No. 2, May, pp. 369–394.

Hobijn, Bart, and David Lagakos, 2005, "Inflation inequality in the United States," *Review of Income and Wealth*, Vol. 51, No. 4, December, pp. 581–606.

__________, 2003, "Social Security and the Consumer Price Index for the elderly," *Current Issues in Economics and Finance,* Federal Reserve Bank of New York, Vol. 9, No. 5, May.

Internal Revenue Service, Statistics of Income, 2010, *Individual Income Tax Returns 2008*, IRS Publication, No. 1304 (Rev. 07-2010), Washington, DC: U.S. Department of the Treasury, available at www.irs.gov/pub/irs-soi/08inalcr.pdf.

McGranahan, Leslie, 2008, "Food inflation and the consumption patterns of U.S. households," *Chicago Fed Letter*, Federal Reserve Bank of Chicago, No. 255, October.

McGranahan, Leslie, and Anna Paulson, 2006, "Constructing the Chicago Fed Income Based Economic Index—Consumer Price Index: Inflation experiences by demographic group: 1983–2005," Federal Reserve Bank of Chicago, working paper, No. WP 2005-20, revised November 2006.

Meyer, Bruce D., and Dan T. Rosenbaum, 2001, "Welfare, the earned income tax credit, and the labor supply of single mothers," *Quarterly Journal of Economics*, Vol. 116, No. 3, August, pp. 1063–1114.

National Commission on Fiscal Responsibility and Reform, 2010, "The moment of truth," report, Washington, DC, December, available at www.fiscalcommission.gov/sites/fiscalcommission.gov/files/documents/TheMomentofTruth12_1_2010.pdf.

Rupp, Kalman, and Charles G. Scott, 1995, "Length of stay on the Supplemental Security Income disability program," *Social Security Bulletin*, Vol. 58, No. 1, Spring, available at www.ssa.gov/policy/docs/ssb/v58n1/v58n1p29.pdf.

Schott, Liz, and Zachary Levinson, 2008, "TANF benefits are low and have not kept pace with inflation," Center on Budget and Policy Priorities, report, November 24, available at www.cbpp.org/pdf/11-24-08tanf.pdf.

U.S. Bureau of Labor Statistics, 2007, "Chapter 17. The Consumer Price Index (Updated 06/2007)," in *BLS Handbook of Methods*, Washington, DC, available at www.bls.gov/opub/hom/pdf/homch17.pdf.

U.S. Census Bureau, 2010, *Statistical Abstract of the United States: 2011*, Washington, DC, available at www.census.gov/compendia/statab/2011/2011edition.html.

U.S. Department of Housing and Urban Development, 2011, *FY 2010 Annual Performance Report*, Washington, DC, March, available at http://portal.hud.gov/hudportal/documents/huddoc?id=fy2010apr.pdf.

White House, Office of Management and Budget, 2011a, *Analytical Perspectives, Budget of the United States Government, Fiscal Year 2012*, Washington, DC: U.S. Government Printing Office, available at www.whitehouse.gov/omb/budget/Analytical_Perspectives.

__________, 2011b, *Historical Tables, Budget of the United States Government, Fiscal Year 2012*, Washington, DC: U.S. Government Printing Office, available at www.whitehouse.gov/omb/budget/Historicals.

Clearing over-the-counter derivatives

Ed Nosal

Introduction and summary

The recent financial crisis highlighted some of the potential problems associated with over-the-counter (OTC) derivatives markets. Prior to the financial crisis of 2008, the over-the-counter market was not required to "clear" transactions. This changed with the signing of the new financial reform legislation, the Dodd–Frank Act, on July 21, 2010. Going forward, most OTC derivatives will be cleared through a particular set of institutional arrangements: a regulated clearinghouse.

The financial crisis exposed some significant cracks in the OTC derivatives markets,[1] as exemplified by the case of insurance company AIG. We now know that AIG sold a gigantic amount of OTC credit default swaps (insurance contracts against defaults), with notional value over $440 billion. Since AIG did not take positions to offset their credit default swap exposure, one might conjecture that this could cause problems for AIG and financial markets.[2] Namely, if the economy were hit by an adverse aggregate shock that either caused a very large number of swap payouts to move against AIG or weakened its balance sheet (or both), then AIG may not have the funds to perform on its OTC obligations. And even if it did have the funds, AIG may choose to default (referred to as a strategic default) on its credit default swap obligations. In either case, AIG's counterparties would not receive the payoffs they were expecting, and this could have a ripple effect on their counterparties and the broader markets.

In fact, during the financial crisis, AIG had to raise money in debt markets in order to make payments on the real-estate-related credit default swaps that moved against it. This new debt issue, along with the large potential future losses that AIG could experience, resulted in AIG's debt being downgraded by the credit-rating agencies. As a result of this downgrade, AIG was required to post billions of dollars of collateral for its existing swap contracts. If the collateral had not been posted, then cross-default clauses in other contracts that AIG had written would have been activated, requiring AIG to settle these contracts immediately. AIG did not have the liquidity on hand to make all these payments, even though its core business activities were in fine shape. Without some help, AIG would have defaulted on its contractual obligations. Hence, the deterioration of the real estate market put significant financial stress on AIG, putting not only AIG at risk, but also AIG's counterparties and their counterparties, and so on. In the end, collateral was posted for AIG's positions thanks to a massive government bailout.[3]

How could this scenario have been avoided? One could argue that clearing AIG's OTC derivatives contracts would have prevented this negative outcome. Intuitively, one can think of clearing as a set of institutional arrangements that are designed to enhance contractual performance. This includes a wide range of procedures that are implemented after a buyer and seller agree to the contract terms and before final settlement occurs. For example, if AIG had been required to clear its credit default swaps, it would have had to set aside collateral when it initially negotiated the contracts. When the swap contracts turned against AIG, it could have used the collateral to satisfy its positions, instead of borrowing, which, in turn, would have prevented the credit downgrade in its debt.

Ed Nosal is a vice president and senior research advisor in the Financial Markets Group of the Economic Research Department at the Federal Reserve Bank of Chicago. *The author would like to thank Robert Steigerwald for many hours of conversation on clearing and for discussing earlier drafts of this article. He would also like to thank Lisa Barrow, Jeff Campbell, Adam Cooper, Randall Costa, Godfried De Vidts, Darrell Duffie, Marcus Katz, Helen Koshy, Adi Li, David Marshall, Bob McDonald, and Ketan Patel for discussions and useful comments on this article.*

Clearing would also have led to greater transparency. Even absent posting collateral, if information regarding AIG's credit default swap portfolio was publicly available, it is likely that AIG's credit default swap portfolio would have been much smaller as counterparties would have been reluctant to enter into arrangements with such a highly leveraged entity. These sorts of remedies have, in fact, been mandated in the recent Dodd–Frank Act, which requires the vast majority of OTC derivatives to be cleared through a regulated clearinghouse.

Since clearing is a not a costless activity, two questions naturally arise: 1) Who should be allowed to participate in a clearinghouse as a clearing member?; and 2) Which derivative contracts should be cleared? A recent public policy symposium held at the Federal Reserve Bank of Chicago addressed these questions.[4] In this article, I review the topic of clearing from an economics perspective and provide the reader with a framework to think about clearing issues. The analysis can offer answers to basic but important questions, such as: What is clearing? Why is it important? What is the role of a clearing member? And why is risk management important for clearing? Second, I use this framework to review some of the ideas that were expressed at the symposium regarding who should be allowed to participate in the clearing process and what should be cleared.

I conclude that: 1) although certain criteria must be met for individuals or institutions to participate in the clearing process, one criterion that should *not* be imposed is that participants must possess massive amounts of capital; and 2) derivative contracts negotiated by "end users," corporations that hold the contracts until they expire, should not be treated any differently from contracts that are negotiated by other firms that do not necessarily hold them until expiration.

In the next section, I provide a simple economic environment to think about the clearing concept. The economic environment presented in this section has the benefit of being transparent, but at the cost of being quite simple. In the two following sections, I complicate the environment slightly, in order to discuss the behavior of OTC participants and risk management—important concepts for clearing. I describe how a clearinghouse might be structured. Then, I present and analyze the overarching themes of the symposium. Finally, I analyze the debate about clearinghouse membership and end-user exemptions.

A simple model of clearing

A natural starting point for any economic investigation is the Arrow–Debreu model. In the model, people are fully aware of all possible future contingencies and are able to write contracts, at the beginning of time, for delivery and acceptance of all possible commodities, where a commodity is distinguished by date, location, and state of the world. Contracts are mediated by markets, and there is a market and price for each commodity, that is, markets are "complete." The model determines one of the most important concepts in economics, which is that of relative prices.

Since people are fully aware of all future contingencies and markets are complete, all decisions regarding how much to buy and sell can be made at the beginning of time. Hence, all trading of contracts occurs at one point in time—at the beginning of time—and as time moves on, people simply make or accept delivery of commodities based on their contracts. Spot markets are not needed at future dates since, as the economy moves in time, people do not learn anything that they did not already know at the beginning of time. If spot markets opened up at a future date and trade occurred, that would imply that people made unexplainable mistakes at the beginning of time, that is, they are irrational. So, in the Arrow–Debreu world, all deals are struck at the beginning of time, and people trust one other to perform as specified in the contingent contracts.

The elegance and simplicity of the Arrow–Debreu model owe a lot to the absence of frictions in the environment. For example, people do not have to find one another or bargain over prices and quantities or worry about contractual performance. All of these frictions are assumed away. In fact, the lack of frictions in the Arrow–Debreu model greatly simplifies the nature of social interaction: There is none! People observe equilibrium prices and trade only against their own budget constraint, which means that the value of what they buy cannot exceed the value of what they sell. The model, however, cannot explain a number of important things, such as how goods are exchanged or why the institutions of money, banking, and clearing exist.

If we want to understand these important institutions, we must introduce some explicit frictions into the benchmark model. For example, if we want to have spot markets open up over time, we can introduce a search friction. If it takes time and effort to find and purchase goods and services, then it will not be possible to contract for all purchases and sales of commodities at the beginning of time. If there are informational frictions, then contracts can only depend upon things that are verifiable, and complete state-contingent contracts are not feasible. If there is a commitment friction, then things like money, banks, and other institutions may arise to help alleviate the commitment problem. And finally, we might also want to take into account legal frictions. These frictions may explain some risk-management practices that would be a puzzle if a lack of commitment was the only friction. Although I use all of the above-mentioned frictions in my analysis, the commitment friction—people cannot commit to undertake future actions—and the information friction play prominent roles. Now, I illustrate the importance of two prominent clearing processes—novation and the posting of collateral.

A farmer plants seeds today that produce wheat tomorrow, and a baker needs wheat tomorrow to bake bread. The price per bushel of wheat tomorrow can take one of two values, say, \$5 or \$15, that are equally likely. The farmer and baker are risk averse, meaning that they prefer to agree today to exchange one bushel of wheat tomorrow for tomorrow's expected price of \$10, as opposed to buying or selling at the spot price tomorrow of \$5 or \$15.

The farmer and baker may be able to get their mutually preferred outcomes if they enter into a forward contract. A forward contract is a particular kind of derivative contract, where the farmer promises to deliver a commodity, one bushel of wheat, tomorrow in exchange for \$10; and the baker promises to deliver \$10 tomorrow in exchange for the commodity. If the farmer and baker can commit to these promises, then they can get their preferred outcomes—wheat for \$10—and that's the end of the story. But if the farmer and baker cannot commit, then delivery and exchange of wheat for \$10 won't happen. To see this, suppose that the price of wheat turns out to be \$5 and the baker does not accept delivery from the farmer and, instead, purchases wheat on the spot market, that is, the baker *strategically defaults* on the agreement because the spot price is lower. This strategy gives the baker an extra \$5, compared with the strategy of performing his contractual obligation. Similarly, the farmer can get a net benefit of \$5 per bushel by strategically defaulting when the price of wheat tomorrow is \$15. Although the farmer and baker would *like* to exchange wheat for \$10 tomorrow, their lack of commitment prevents this from happening. If the farmer and baker could somehow bind themselves to a \$10 per bushel agreement, then they would do so, so long as the cost of binding isn't too great. This is where the notion of clearing comes in.

One way the baker and farmer may be able to bind themselves to the contract is for each of them to provide \$5 of collateral upfront. The \$5 of collateral is used to cover any losses incurred by a counterparty should the other counterparty fail to perform on the contract. For example, if the price of wheat is \$5 and the baker reneges on the contract, then he loses his \$5 of collateral, which is given to the farmer. In effect, the baker pays \$10 for the wheat. It would seem that with the introduction of collateral, the parties should not have an incentive to renege on their contracts. Maybe.

One tricky issue is who or what is to hold the collateral? Notice that the lack of commitment cannot be overcome by simply having each party hold the other's collateral. To see this, suppose the spot price turns out to be \$15. In this situation, the farmer does best for himself by selling his wheat on the spot market and keeping the baker's collateral (and the baker keeps the farmer's collateral). Here, one holding of collateral simply offsets the other one and does not guarantee performance. It appears that a third party is needed to hold the collateral of the baker and farmer.

With the introduction of a third party, things could work as follows. If the baker and farmer perform on their contract, then the third party returns the collateral to each of them. If, however, one party defaults and, as a result, harms the other party, then the third party can use the collateral of the nonperforming party to compensate the other party for his losses. For example, if the farmer defaults when the price of wheat is \$15, then the third party gives the baker the farmer's collateral, as well as his own. From the farmer's point of view, he will pay \$10 for wheat if he chooses to default, so he now has no incentive to default. It appears that the introduction of a third party that holds the collateral of the farmer and baker implies that they will each perform their contractual obligations. Maybe.

In order for the three-party arrangement to work, it is necessary for the third party to be able to verify which party reneges in the event of contractual nonperformance. Suppose that there is an informational friction. In the last example, where the spot price is \$15, the farmer can renege and claim to the third party that he attempted to deliver the wheat to the baker but, for some reason, the baker refused to take delivery. The farmer, then, could argue because of the baker's nonperformance, he had no choice but to sell his wheat on the market and he should not forfeit his collateral. Hence, if the third party cannot perfectly observe or verify the actions of the farmer and baker, then it will be unable to determine which party in fact reneged. This implies that a simple third-party mechanism—one that simply holds collateral—cannot guarantee performance.

One way around this verifiability problem is to have all the transactions related to the contract go through the third party. That is, instead of having the farmer deliver wheat to the baker and the baker deliver \$10 to the farmer, all deliveries are made to the third party. The third party then "redelivers" the wheat and money according to the original contract. Under this scheme, the initial contractual obligation between the baker and farmer is discharged and replaced with two new contracts: one between the third party and the baker and another between the third party and the farmer. This process, called *novation*, makes the third party a *central counterparty*, or CCP, to the original parties of the contract. Notice that novation circumvents the verifiability problem: Since the farmer and baker make deliveries to the CCP, the CCP is able to determine whether each party has performed its obligations. The CCP is now able to transfer the collateral it holds to the appropriate party in the event of nonperformance.

The above example illustrates how novation and collateral can substitute for commitment. Collateral and novation are only two possible ingredients or processes that fall under the rubric of clearing. Things like trade matching and confirmation (that is, checking to make sure party A entered into a contract and that party B is, in fact, on the other side of the contract), information warehousing (that is, compiling in an accessible manner the set of all the trades that have taken place), and risk management are also part of the clearing process.

In the example, the commitment friction is needed for a CCP to emerge; without it counterparties would not strategically default and there would be no need for a CCP. In practice, one might think that even if counterparties cannot commit, concern about their reputations might provide sufficient inducement for performance in the absence of collateral and CCPs. For example, a major financial player will not strategically default on a contract (which is not backed by collateral) for only a small gain today, since that would damage its reputation as a reliable institution in the future. There is a cost associated with being viewed as unreliable: Potential future counterparties will choose not to enter into mutually beneficial contracts with an "unreliable" institution. (These counterparties fear the institution will not perform on its contract.) If a counterparty's concern for its reputation is sufficiently strong, then it will perform on its contracts even though it lacks commitment. If this is always the case, then my claim that a commitment friction is responsible for the emergence of CCPs seems inappropriate. That is, reputation will enforce performance and a CCP is not needed. I now examine this issue.

Commitment and reputation

An implication of the commitment friction, as illustrated in the farmer–baker example, is that if counterparties do not post collateral with the CCP, then one of the counterparties will always strategically default. Yet, in practice, we routinely observe people and organizations honoring contracts that are not backed by collateral. We also observe contractual performance when collateral is posted on a bilateral basis, that is, a third-party CCP does not hold the collateral. Do these observations imply that CCPs and other clearing processes arise for reasons other than a commitment friction? One could argue that counterparties perform on their contracts because they care about their reputations, and concern for one's reputation outweighs the commitment friction. Below, I argue that a commitment friction is still relevant even when parties care about their reputations. Hence, we should take the commitment friction seriously when thinking about clearing.

Here, I modify the farmer–baker example so that I can discuss the notion of reputation. Instead of having a "one-shot" contracting relationship between the farmer and baker, suppose that the farmer and baker repeatedly interact with one another. The repeated nature of the farmer–baker relationship may provide an incentive for them to perform their contractual obligations even when there is a commitment friction and no collateral is posted. Why? Because performance today implies that parties will choose to contract with each other in the future, and this interaction is beneficial to both parties. Or, put another way, a party will only enter into a new contract with a counterparty that has a reputation for performing.

To see this, suppose the farmer and baker agree to exchange wheat for $10. Suppose further that if the farmer fails to deliver wheat when the spot price is $15, then both parties decide that they will never enter into another contractual relationship with each other. The baker believes that if they do, then the farmer will always fail to deliver when the price is $15. With this belief, the baker has no incentive to enter into a contract with the farmer. So, in contrast to the original example, where the interaction between the farmer and baker was a one-shot affair, there is now a cost associated with not performing, which is the inability for the farmer and baker to stabilize the price of wheat at $10 in the future. In deciding whether or not to perform when the price of wheat is $15, the farmer compares the immediate benefit associated with nonperformance, which is $5, with the future benefit associated with performance, which is the ability to sell wheat at $10 in future periods. If the latter exceeds the former, the farmer will perform; otherwise he will strategically default.

One might expect in "normal" times that counterparties will perform their contractual obligations even when no collateral is posted. This is because the benefit associated with future contracting outweighs the short-term gain associated with default. However, in times of "stress," which can be characterized by either large price movements and/or weakened balance sheets, the benefit associated with reneging can look quite attractive.

For example, if there is a huge price movement, a counterparty may find it in its best interest to avoid taking a large loss today, even though it will be unable to enter into a contract in the future. (And this may be the case even when the reneging counterparty has a strong balance sheet and the resources to perform.) Alternatively, if a counterparty has a weak balance sheet, fulfilling a contractual obligation may make the firm insolvent or vulnerable, whereas reneging will allow it to "live" for another day or become stronger. Faced with these choices, the counterparty may choose to strategically default. It is important to emphasize that during stressful times, parties that anticipate receiving positive contractual payments are really counting on their counterparties to perform. Yet, it is precisely at such times that a counterparty is most likely to (strategically) default.[5] If, from a social perspective, it is important to mitigate contractual defaults during times of stress, then we know from the earlier example that performance can be enhanced if collateral is posted with the CCP. Hence, even though parties will perform on their contractual obligations in the absence of posting collateral in most circumstances, a CCP will still emerge and require counterparties to post collateral. The reason is that performance is particularly important to counterparties (and society) during times of stress, but this is a time when their counterparties may not care about their reputations and may strategically default.[6]

The above discussion implies that when there is a commitment friction, a CCP may emerge to enhance performance even though counterparties care about their reputations.

Risk management

In practice, CCPs devote considerable resources to risk management. Unfortunately, the example I introduced earlier is too stylized to illustrate the notion of risk management because the CCP does not face any risk. In the example, the size of the CCP's collateral holdings eliminates the risk that it will fail to perform in the event that one of the counterparties fails to perform. The CCP holds \$10: \$5 from the farmer and \$5 from the baker. In the event that either the farmer or baker defaults, the CCP needs \$5 to guarantee performance, which it has. In the real world, however, prices of goods or assets do not move in nice, finite, discrete amounts over a specific time interval. Although short-term price movements are typically not that large in relative terms, one can't rule out enormous price movements over a period of a few hours. One can imagine modifying the example to allow for a continuous distribution of spot prices characterized by low probabilities of large price movements. In that case, the CCP can only guarantee performance if it holds enormous amounts of collateral. But this particular solution to the performance problem may not be an attractive one. Posting collateral is costly for counterparties, as they have more productive uses for their resources. If a CCP demands huge amounts of collateral—so as to guarantee performance for any conceivable price movement—then the farmer and baker may simply stop using forward contracts. In this situation, forward contracts would become extremely costly to use. The farmer and baker may therefore prefer to transact on spot markets and face the (lower cost) price risk. This is where risk management comes in.

The CCP can economize on collateral by "guaranteeing" performance with a high probability (but less than one). So, although there is a possibility that the CCP will default on its obligations, it minimizes this possibility by managing the risk of failure that it faces. A CCP can do this by, for example, requiring adjustments to collateral when it perceives that risk has changed and making provisions for additional resources should its collateral holdings prove to be insufficient to cover a default by one of its counterparties. I will now develop these ideas a bit further.

The amount of collateral that counterparties post can depend on a number of things. It can depend on the volatility of the price of the commodity or assets that underlie the derivative contract.[7] There is a positive relationship between the volatility of the price of the underlying asset and the volatility of the value of the derivative contract. If the underlying asset has low price volatility, then it will require less collateral than an underlying asset with high price volatility. The idea here is that if a counterparty fails, then, on average, a low-price-volatility underlying asset will require a smaller amount of the CCP's resources to ensure contractual performance than a high-price-volatility underlying asset. Also, if the price volatility of the underlying asset changes over time, then so should the amount of collateral that counterparties post with the CCP. For example, if the volatility decreases, the CCP will transfer some of the collateral back to the counterparty; if it increases, then it will require additional collateral from the counterparty.

For longer-lived derivative securities, the value of a particular position will change over time. For example, if the price of the underlying asset, say, wheat, in a forward contract increases over time, then the value of the forward contract for the counterparty who delivers it, the farmer, falls, since he is delivering an asset whose value is much greater than the delivery price. If this counterparty fails to deliver, then the CCP faces a very high performance cost. The CCP can eliminate these "accumulated liabilities" by requiring counterparties to settle these liabilities on a periodic basis. That is, the CCP marks-to-market its contracts—directly passing resources from counterparties whose contracts have lost value to those whose contracts have gained value.

The amount of collateral required by the CCP can also depend upon the "liquidity" of the contracts it novates. In practice, if a counterparty defaults, the CCP sells the defaulting counterparty's positions. If a position is "liquid," it can be sold quickly and at a low cost. For example, contracts that are exchange traded and cleared, such as futures contracts, are very liquid.[8] The market provides a fairly accurate estimate of the value of a position, and even a very large position can be sold to the market over a relatively short period. On the other hand, specialized OTC derivative contracts are much less liquid because their estimated value is subject to great variation and they are traded on a bilateral basis. Hence, it may be difficult to sell a large position of OTC contracts on short notice at a price that is at or near their estimated value. Since less liquid contracts are more costly to trade than liquid ones, a CCP will require higher collateral for OTC contracts.

The CCP may have access to resources other than collateral to help it facilitate contractual performance in the event of a counterparty failure. The CCP may require all of its members, that is, members that have their contracts novated by the CCP, to contribute to a guarantee fund. Members make this contribution when they join the CCP, before the CCP novates any of their contracts. This fund could be accessed in the event a member defaults and its collateral is insufficient to guarantee performance. (This happens when the value of the defaulting member's portfolio—which includes current payments—plus the value of the member's collateral is negative.) Such an arrangement is sometimes called loss mutualization, because losses are shared or *mutualized* among the nondefaulting or surviving members. If the defaulting member's collateral and the entire guarantee fund are insufficient for the CCP to perform, then the CCP members' agreement may require them to provide additional resources to ensure performance.

Except for the requirement that members provide additional resources in the event of a member default, the risk-management strategies described above are consistent with the commitment and informational frictions. That is, contributions to the guarantee fund, the posting of collateral, and marking-to-market represent various payments that members make and receive that are designed to enhance commitment. Importantly, these payments are made before any default occurs. The requirement to provide additional resources in the event of a member's default is subject to a commitment friction. Since the additional payment to the CCP occurs after a default, members may choose to honor their promises or not. In particular, in times of stress, members may choose not to honor their promises because the benefit associated with being a CCP member in the future is less than the resources they have to sacrifice today. The analysis regarding whether members will contribute extra resources to a CCP after a member defaults is the same as in the farmer–baker example in the previous section, "Commitment and reputation."

To summarize, risk management is an important element of the clearing process because the CCP's performance guarantee is only as good as its risk-management strategy. Risk-management strategies, such as marking-to-market, making contributions to a guarantee fund, and adjusting collateral holdings when perceived risk changes, are consistent with the commitment friction.

The structure of central counterparties

In practice, a CCP has a set of members. Only members can clear contracts with the CCP. That is, a CCP novates only those contacts that are presented by its members. Counterparties that are not members of a CCP—let's call them customers—have their contracts novated by a CCP member. In this arrangement, the CCP guarantees contractual performance for its members, and a CCP member guarantees contractual performance for its customers. Hence, there is a tiered, but separated, relationship between customers, CCP members, and the CCP. There is no direct contractual relationship between the CCP and customers.

If a clearing member defaults on its contractual obligations with the CCP, the CCP guarantees performance of the defaulting member's contracts. That is, the CCP will perform its contractual obligations for all of its nondefaulting or surviving CCP members. As a result, all surviving CCP members will be able to perform their contractual obligations with their customers and the CCP. It must be pointed out, however, that the CCP guarantee does not (typically) extend to customers of a defaulting CCP member. Since customers' contracts are guaranteed by CCP members, they effectively lose any performance guarantee when the entity that guaranteed performance for them—the defaulting member—no longer exists.[9]

Just as in the farmer–baker example, CCP members post collateral with the CCP, and customers post collateral with CCP members. So, the collateral that CCP members post with the CCP can come from their customers or from the members directly for contracts that they entered into on their own account.

A clearing member is responsible for the performance of the contracts that it brings to the CCP. So, if a customer defaults on its contractual obligation to a

CCP member, the member must step in and ensure performance or be in default with the CCP. Effectively, the defaulting customer's contracts become the contracts of the member. The CCP member, however, does not have to hold onto the contracts associated with the defaulting customer as part of its portfolio; the member can always sell them. In either case, the CCP member will receive the collateral that was posted for the defaulting customer's position. If the collateral requirements were appropriately calculated, then they should cover both the payments made by the member and any losses associated with either holding or selling the defaulting customer's position. If, for some reason, the collateral is insufficient to cover the losses, the CCP member must absorb the losses or be in default to the CCP. The member will typically be willing to absorb these losses for the same reason as the farmer is willing to deliver wheat to the baker at $10 when the spot price is $15 in the absence of collateral. The member values the future benefit associated with being a CCP member more than the short-term benefit of walking away from the losses. In times of stress, however, a CCP member may choose to default on its performance obligations.

Debate about central counterparty membership

At the heart of the debate over CCP membership criteria is the liquidity of OTC contracts. The debate can be loosely characterized as follows. One side believes that because OTC contracts are not very liquid, a CCP member must be able to assume the portfolio of a defaulting member. This necessarily implies that members must have significant capital (typically at least $5 billion) available for clearing purposes. The other side believes that the illiquidity of OTC derivatives is overstated; for example, OTC interest rate derivatives and credit default swaps are quite liquid. If a CCP member defaults, then there are methods available to dispose of the portfolio, other than requiring another member to purchase it. And finally, if a requirement for membership is significant financial resources, then membership will be limited to a very small set of financial institutions, which in turn could give these institutions undue market power. I examine these views, starting at the heart of the debate: liquidity.

Contracts are said to be liquid if the value, or "fair price," of the contract can be accurately determined on an ongoing basis and large amounts of contracts can be bought or sold at or near the fair price in a short period.

A CCP adopts risk-management strategies to enhance its performance guarantee. These strategies—such as collateralizing positions, marking-to-market, and disposing of a defaulting member's portfolio—are easier to implement when contracts are liquid. The amount of collateral posted depends, in part, on the volatility of the price of the underlying asset and of the value of the OTC contract. Higher volatility implies that more collateral should be posted, and changes in volatility imply that the amount of collateral posted should also change. When contracts are liquid, good estimates for levels of volatility can be obtained. The CCP can be reasonably assured that the process of periodically marking-to-market members' positions will not leave the CCP with additional liabilities in the event of a member's default when contracts are liquid. Since large positions can be sold quickly at or near the fair price when contracts are liquid, the CCP will be able to efficiently dispose of a defaulting member's portfolio. A final risk-management strategy, contributions to a guarantee fund, provides an additional buffer for the CCP against losses from a member's default. The size of a member's contribution determines the maximum notional value of contracts (or risk) that the member can bring to the CCP.

If all contracts were liquid, there would be no debate associated with CCP membership. Any counterparty that could post the required collateral for its positions and contribute to a guarantee fund would be able to become a member.

Because some contracts are not very liquid, however, any sensible risk-management strategy, *independent of the structure of membership*, will require higher levels of collateral from CCP members.

An important aspect of guaranteeing performance is the CCP's ability to sell a defaulting member's portfolio quickly, at or near the fair price.[10] When contracts are highly liquid, the CCP does this by simply selling the portfolio to the market. What can the CCP do when the contracts are not very liquid and there is no market to sell to? One side of the debate over CCP membership says the solution is to require each CCP member to assume, that is, purchase, part of the defaulting member's portfolio. Since the notional value of a defaulting member's portfolio may be quite large, a CCP member will require substantial capital to assume its share. Hence, this solution would restrict CCP membership to those who possess significant capital that can be used for clearing purposes.

There may be problems associated with this solution. Because of the significant capital requirement for membership, the number of individuals or institutions that can qualify for membership will be small. This implies that the amount of liquidity that can be brought into the clearing process is limited (by the wealth of the small number of clearing members). Perhaps more importantly, if there is only a small number of clearing members, then they could use the resulting market power to adversely influence the pricing of clearing services and the pricing of the OTC contracts themselves.

If a clearing member defaults, one side of the debate advocates that the CCP auction the defaulting member's portfolio among a small number of institutions, the CCP members. Auctions share many of the desirable properties associated with exchanges or markets, such as price discovery and a place that brings buyers together with a seller. In other words, an auction can provide (some) liquidity for the objects that are being sold. But it is neither clear nor obvious why the auction would operate more efficiently from society's point of view if it is restricted to only "wealthy" bidders. Since a CCP also serves as an information warehouse—collecting and disseminating the prices of cleared contracts—this information could be used by *anyone*, that is, a CCP member or nonmember, who would like to bid on part or all of the defaulting member's portfolio. Opening up the auction to nonmembers would make more liquidity available to the clearing process. (Of course, if a nonmember purchases part of the portfolio, then those contracts would have to be cleared through a CCP member.) There does not appear to be a rationale to limit the sale of a defaulting member's portfolio to only (wealthy) clearing members.

Another problem created by the wealth restriction for membership lies in the pricing of products. There is always a public policy concern regarding the pricing and supply of services when the number of service providers—in our case, CCP members—is small. That is, prices will be too high and quantities will be too low. Perhaps a bigger potential problem is that since CCP members are free to choose their customers, they may choose to clear only those OTC contracts for which they are direct counterparties, thereby limiting competition in the OTC derivative markets.[11] Hence, the wealth restrictions for CCP membership, which are motivated by clearing considerations, can have adverse affects on the pricing of the OTC derivative contracts. Ironically, the wealth restriction could ultimately prevent these contracts from become more liquid.

The wealth restriction for membership seems artificial. As long as an institution can cover the risk that it brings into the CCP, by providing appropriate levels of collateral and making contributions to the guarantee fund, there does not appear to be any reason to exclude it from membership. In an *unrestricted* membership environment, CCP members would compete for customers by appropriately pricing their services. With this structure, there would not be any obstacles to clearing all "clearable" contracts (not just those for which a member is a counterparty) or moving current OTC contracts onto exchanges. Exchange trading would improve both the liquidity of the contracts and the CCP's performance.

In summary, restricting CCP membership limits the amount of liquidity in the clearing process. In addition, membership restrictions can have adverse affects on the provision and pricing of both clearing services and OTC contracts.

End-user exemptions

"Since we weren't part of the problem, we shouldn't have to pay." This statement nicely summarizes the sentiment of many nonfinancial corporate end-users of OTC derivative products regarding that part of the Dodd–Frank legislation that mandates clearing for most OTC derivative contracts.

In most cases, nonfinancial corporations purchase OTC derivative contracts to hedge their business risks. These contracts are attractive because, unlike exchange-traded derivative contracts, they can be tailored to the firm's business needs (for example, in terms of timing of payments). Because these firms are using the contracts for hedging rather than speculative purposes, they usually hold onto them until they expire.

Currently, if a nonfinancial corporation wants to purchase an OTC derivative, such as a swap, it negotiates the terms directly with a dealer. The dealer typically does not require the nonfinancial firm to post collateral. However, the firm does pay a premium over similar products that require the buyer to post collateral. Nonfinancial firms claim that a requirement that all OTC contracts be centrally cleared will raise their cost of hedging, because they will now be required to post collateral. These firms argue that since they use the contracts to hedge their business risks and, by and large, they did not default on their derivative contracts during the financial crisis, they should not have to bear this cost.[12]

I conclude that, even in the best case scenario, the cost of hedging for nonfinancial corporate end-users will increase when their contracts are cleared. In the best-case scenario, the dealer posts collateral and the end-user does not. This is the best-case scenario from an end-user's perspective, because it does not have to post collateral. I now provide the details for this scenario. Suppose that the dealer that negotiates the corporate end-user's swap has a large diversified portfolio of end-users. As well, assume that corporate end-users will never (strategically) default on their swap obligations. That is, whenever end-users make the calculation, the benefit associated with continued access to the swap market exceeds the benefit of defaulting on a swap payment. Hence, the only time an end-user defaults on a swap payment is when it is insolvent (or bankrupt). If one assumes that the probability of a firm becoming insolvent is independently distributed across firms,[13] then the dealer understands that a certain proportion of its swap contracts will end up in default. The dealer can charge a premium that reflects the proportion of swap contracts that will fail. From the end-user's perspective, this can be interpreted as an insurance premium. A default by an end-user will not create any problems for the broker-dealer since defaults are anticipated and priced. In this ideal world—where end-users never have an incentive to strategically default and (nonstrategic) defaults are uncorrelated—end-users do not have to post any collateral. Note that this result is consistent with my "simple" model of clearing, where collateral was required to guarantee performance because either the farmer or baker would always have an incentive to strategically default.

Prior to the recent financial crisis, the dealer typically would not have been required to post collateral for its positions with the end-user. Most of the time, dealers would contractually perform because it was in their best interests to do so. However, in times of stress, the dealer might default on its swap obligations. As a result, the end-user would not receive any payment—which may be critical during times of stress—and would be one of many creditors seeking remedy from the dealer. Clearly, contractual performance would be enhanced if the OTC contracts were cleared by having dealers post collateral for their positions with corporate end-users.

But posting collateral is not cheap. When dealers post collateral for their swap contracts with corporate end-users (as mandated by the Dodd–Frank legislation), they will pass some proportion—possibly all—of the associated cost to the end-users. Therefore, even if corporate end-users do not post collateral for their positions, their cost of hedging will increase.

Up to this point, I have assumed that corporate end-users don't strategically default and that defaults by nonfinancial firms are uncorrelated. These assumptions are unrealistic. In times of stress, when there are large price movements and/or weakened balance sheets, end-users may find it in their best interests to strategically default on a swap payment, even though they have the resources to pay. If many end-users default, then the insurance premium that dealers charged them will be insufficient to cover the dealers' losses. If, from a social perspective, it is important that dealers do not experience large-scale loses on their swap contracts during times of stress, then end-users' swap positions should be properly cleared through a CCP, which will require them to post collateral.

Although nonfinancial firms, by and large, purchase OTC contracts for the purpose of hedging, it is not at all obvious that these entities do not pose a threat to the stability of the financial system. Nonfinancial corporate end-users represent a relatively large share of the OTC market, 10 percent to 15 percent. If these firms receive a correlated shock that weakens their ability to perform, they may transmit this adverse shock to the balance sheets of the dealers. The potential effects of this shock can be greatly mitigated by requiring them to post collateral for their positions.

Conclusion

In the first part of this article, I sketched out a framework for thinking about clearing. I used the insights from this framework to examine two prominent themes from a recent symposium on clearing. In terms of CCP membership, there is an alternative (unrestricted) CCP structure that is at least as effective as one that requires members to have substantial capital. The alternative structure has the added benefit (which could be huge) of promoting both competition and the provision of liquidity in clearing and in the OTC derivatives markets. In terms of end-user exemptions, the cost of hedging will increase for end-users even if they are not required to post collateral. Since end-users' positions are nontrivial, in the sense that (correlated) defaults by end-users can weaken the ability of their dealers to perform, they should post collateral in order to strengthen dealers' ability to perform.

NOTES

[1]Derivatives are financial contracts whose value is linked to the price of an underlying commodity, asset, rate, index, or the occurrence or magnitude of an event. These contracts are traded both on traditional exchanges and over the counter.

[2]In 2008, the notional value of all AIG's derivative contracts, including credit default swaps, was as high as $2.7 trillion.

[3]See www.freakonomicsmedia.com/2008/09/18/diamond-and-kashyap-on-the-recent-financial-upheavals/ for a helpful Q&A about the financial market disruption in late 2008.

[4]The symposium agenda is available at www.chicagofed.org/webpages/events/2010/public_policy_symposium_on_OTC_derivatives_clearing.cfm.

[5]Some people argue that large companies do not strategically default because the reputational cost is too great. So, if these companies do default, it is because they are either bankrupt or insolvent. But, if a company senses it may become insolvent, it may attempt to counteract this by taking defensive actions such as rescheduling debt and other payments in an attempt to save itself. One can interpret this as strategically defaulting—the company has the resources to pay current bills, but chooses to withhold these payments.

[6]We observe this in the context of sovereign debt. For example, a country has the resources to pay for debt, but chooses not to because the current costs of doing so are too high, say, because the country would have to increase taxes on its citizens when the economy is weak. This is a costly decision—when a country repudiates debt payments, its reputation will take a hit, in the sense that it will be shunned in international debt markets in the foreseeable future.

[7]In the earlier example, the derivative contract is a forward contract and the commodity that underlies the derivative contract is wheat.

[8]A measure of liquidity for exchange-traded contracts is the bid–ask spread. A low bid–ask spread indicates that the contract is liquid; a high bid–ask spread indicates that it is not very liquid.

[9]If the customers' collateral payments are segregated from the member's own collateral and the member defaults on its own positions, then the customers' collateral is protected. In practice, customers of the defaulting member would transfer their accounts to one or more nondefaulting members, who then can "novate" the customer contracts of the defaulting member.

[10]The CCP may be able to avoid receiving a low selling price if additional time is taken to sell the position. There are two possible problems, however. First, there is a chance that the price of the contracts can move against the CCP during a protracted selling period. This implies that the CCP would require additional collateral from its members. Second, the CCP may not have the luxury of time on its side to dispose of the portfolio.

[11]If a clearing member buys or sells a contract from customer A, then I say that customer A is a direct counterparty of the clearing member. If customer A buys or sells a contract from customer B, a clearing member can limit competition by choosing not to novate the contracts.

[12]However, there have been defaults by nonfinancial firms that disrupted the broader markets—for example, Enron.

[13]This assumption is almost certainly false, that is, insolvencies will be correlated. I make this assumption because it provides the best-case scenario for the corporate end-user. Later, I discuss the implications of relaxing this assumption.

Worker flows and matching efficiency

Marcelo Veracierto

Introduction and summary

One of the best known facts about labor market dynamics in the U.S. economy is that unemployment and vacancies are strongly negatively correlated, an empirical relationship called the "Beveridge curve."[1] In recent times, however, large deviations from the Beveridge curve have been observed. In particular, vacancies have increased quite significantly since mid-2009, but this phenomenon has not been accompanied by a substantial decrease in the unemployment rate (see figure 1). This failure of the Beveridge curve has surprised many economists and has been interpreted as evidence of mismatch, that is, of increased frictions in the process through which workers meet job opportunities (for example, Kocherlakota, 2010). The purpose of this article is to provide a measure of mismatch in U.S. labor markets and to assess its importance in determining the behavior of the unemployment rate and other labor market outcomes since the start of the latest recession. The framework that I use is a simplified version of the Mortensen and Pissarides (1994) model. Since my purpose is to use it as a tool for organizing and interpreting data, I will abstract from any explicit decision making and focus on the essential structure of the model.

The basic structure of the Mortensen and Pissarides model has three main components. First, it has an aggregate matching function that summarizes the process through which unemployed workers and employers with open vacancies search for each other and meet. It functions very much like a standard production function, with unemployed workers and vacancies entering as inputs of production and the number of matches formed appearing as output. The second element is a free-entry condition for the creation of vacancies. In particular, it is assumed that there is a fixed cost to post a vacancy and that employers create vacancies up to the point at which the expected discounted value of a filled job equals this fixed cost. The expected value of a job is given by the probability of filling the job, which is determined by the aggregate matching function, and by the value of a job. In a full-blown version of the model, the value of a job is endogenously determined by the expected revenues that the job will generate and by the bargaining power of the worker. However, in the simplified version considered in this article, I am silent about the explicit process through which the value of a job is determined. The third main component is a simple accounting relationship that states that the total flows in and out of each labor market state must be equal. A standard approach in the literature is to allow only for two labor market states (employment and unemployment) and to assume that the model is always at its steady state (that is, its long-run equilibrium). However, I consider more flexible specifications in this article.

I use this simple version of the Mortensen and Pissarides model to measure mismatch and evaluate its consequences during the post-2007 recession period. This is not the first article to do this. Two closely related papers are Barlevy (2011) and Barnichon and Figura (2010). Barlevy follows the standard approach by postulating two labor market states, assuming a constant separation rate (that is, the rate at which workers transit from employment into unemployment), and by assuming that the model is always at its long-run steady state. On the contrary, Barnichon and Figura incorporate a third labor market state (nonparticipation) and allow the transition rates between the three labor market states to vary over time. However, similar to Barlevy, Barnichon and Figura assume that the model is always at its steady state and that only unemployed workers enter the matching function.[2]

Marcelo Veracierto is a senior economist in the Economic Research Department of the Federal Reserve Bank of Chicago. *The author is grateful for comments from Gadi Barlevy, Andy Neumeyer, and Juan Pablo Nicolini, as well as seminar participants at the Federal Reserve Bank of Chicago and the Universidad Torcuato Di Tella. He also thanks Hao Zou for research assistance.*

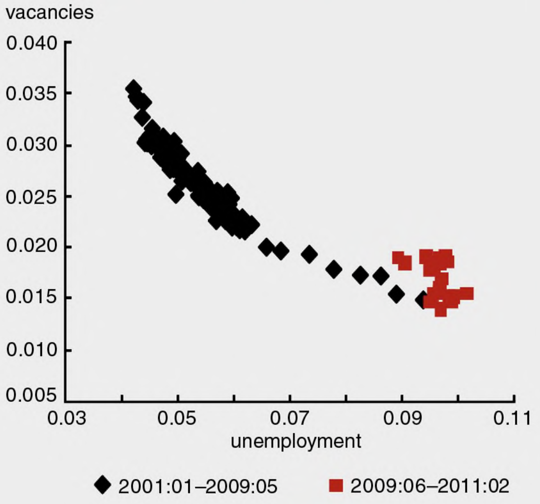
Given the different assumptions made in the literature, I use my model to evaluate how sensitive the results are to the different specifications. I consider the following dimensions. First, I assess the importance of allowing the separation rate to vary over time instead of assuming it to be constant. Second, I evaluate the consequences of specifying three labor market states instead of two. Third, I assess the consequences of assuming that the model is always at its steady state instead of allowing for transitionary dynamics. Fourth, I evaluate the consequences of allowing nonparticipants to enter the matching function instead of assuming that the matching function solely applies to unemployed workers.

I find that the results are extremely sensitive to the alternative specifications. However, in the preferred scenario (which has three labor market states, variable transition rates, transitionary dynamics, and nonparticipants entering the matching function), I obtain the following findings.[3] First, the matching efficiency has been quite volatile throughout the whole sample period (2001:1[January]–2011:2[February]). Second, the matching efficiency has been drifting down since the start of the last recession. Third, the value of filled jobs plummeted between 2007:12 (the start of the latest recession) and 2009:6, but it has recovered quite significantly since then. Fourth, conditional on the observed paths for the value of a job and all transition rates, the drop in matching efficiency since the start of the recession has had only a moderate impact on the unemployment rate: The current unemployment rate would be 1 percentage point lower if the matching efficiency had stayed unchanged. Fifth, the bulk of the increase in the unemployment rate since the start of the recession is accounted for by changes in the transition rates across labor market states. Sixth, the matching efficiency, the value of a job, the transition rates, and the search intensity of nonparticipants all have significant effects on the dynamics of nonparticipation. Since they deemphasize the importance of matching inefficiencies in explaining the large increase in the unemployment rate since the start of the last recession, the results in this paper are consistent with a greater role for policy in achieving improvements in labor market conditions.

In the next section, I consider the case of two labor market states. In the following section, I consider the case of three labor market states and a matching function with only unemployed workers. Then, I consider the case of three labor market states but allow for nonparticipants to enter the matching function. Readers solely interested in learning about the relative contributions to unemployment dynamics should jump to the last section of the paper, which uses the preferred scenario. The first two sections report results using alternative but less satisfactory methodologies.

FIGURE 1

Unemployment rate vs. vacancies



Note: Vacancies and unemployment are normalized by total labor force.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

The case of two labor market states

There are two types of agents: firms and workers. Each firm has one job available, which can either be filled or vacant. The expected discounted value of profits generated by a filled job is equal to J_t units of the numeraire. Posting a vacant job requires k units of the numeraire. There is an infinite number of potential firms. Workers can be in either of two states: employed or unemployed. Employed workers get separated from their current jobs with probability λ_t. Unemployed workers and posted vacancies determine the total number of new matches that are formed according to the following matching function:

$$1)\quad M_t = A_t U_t^{\alpha} V_t^{1-\alpha},$$

where M_t is the total number of new matches, U_t is the total number of unemployed workers, V_t is the total number of posted vacancies, A_t is the productivity of the matching function, and $0 < \alpha < 1$.

Normalizing to one the total number of workers in the economy, the evolution of unemployment over time can be described by the following equation:

2) $U_{t+1} = U_t - M_t + (1 - U_t)\lambda_t$.

That is, the total number of workers that will be unemployed tomorrow U_{t+1} is equal to the total number of currently unemployed workers that do not find a match $U_t - M_t$, plus the total number of currently employed workers that get separated from their jobs $(1 - U_t)\lambda_t$.

Since firms are profit maximizers, the following free-entry condition must be satisfied:

3) $$k = \left(\frac{M_t}{V_t}\right) J_t.$$

That is, the cost of posting a vacancy k must be equal to the probability of filling a vacancy M_t/V_t times the expected discounted value of profits generated by a filled job J_t. If this condition was not satisfied, the total number of vacancies created would be either zero or infinity, depending on the direction of the resulting inequality.

Observe that the productivity of the matching function A_t, the separation rate λ_t, and the expected discounted profits generated by a filled job J_t are exogenous to the model. Given the total number of workers unemployed at date zero U_0, the model generates an endogenous path for $\{M_t, V_t, U_{t+1}\}_{t=0}^{\infty}$.

Steady state

Assuming a constant matching productivity A, a constant separation rate λ, and constant expected discounted profits generated by a filled job J, a steady state of the model economy can be defined as an initial unemployment level $U_0 = U$, such that the endogenous path for $\{M_t, V_t, U_{t+1}\}_{t=0}^{\infty}$ that the model generates is constant over time. That is, that $M_t = M$, $V_t = V$, and $U_t = U$ for every $t \geq 0$. A steady state (M, V, U) can be interpreted as the total matches, vacancies, and unemployment that the economy will converge to in the long run.

From equations 1–3, we have that the conditions a steady state must satisfy are the following:

4) $$M = AU^{\alpha}V^{1-\alpha},$$

5) $$M = (1 - U)\lambda,$$

6) $$k = \left(\frac{M}{V}\right) J.$$

Substituting equation 5 in equations 4 and 6 gives the following simplified steady-state conditions:

7) $$U = \frac{\lambda}{\lambda + A\left(\frac{V}{U}\right)^{1-\alpha}},$$

8) $$k = A\left(\frac{U}{V}\right)^{\alpha} J.$$

Equation 7 defines a negative relationship between unemployment and vacancies and, for this reason, is called the "Beveridge" curve. Equation 8 defines a positive linear relationship between unemployment and vacancies and, since it is defined by a free-entry condition to the posting of vacancies, it is called the "job creation" curve. The Beveridge and job creation curves are depicted in figure 2. The intersection of these curves determines the steady state (U^*, V^*).

It is particularly important to determine what causes shifts in each of these two curves. It is possible to show that an increase in the separation rate λ shifts the Beveridge curve up, an increase in the expected discounted profits from a filled job J does not affect the Beveridge curve, and an increase in the matching efficiency parameter A shifts the Beveridge curve down. In turn, the separation rate λ has no effect on the job creation curve, but an increase in either J or A rotates the job creation curve clockwise. Given these shifts in the Beveridge and job creation curves, we can now determine how changes in λ, J, and A affect the steady-state pair (U^*, V^*). In particular, we can conclude that that an increase in λ increases both vacancies V and unemployment U, that an increase in J increases V and reduces U, and that an increase in A reduces V. The effects of an increase in A on U are unclear from the figure, but substituting equation 8 in equation 7 gives that

$$U = \frac{\lambda}{\lambda + A^{\frac{1}{\alpha}}\left(\frac{J}{k}\right)^{\frac{1-\alpha}{\alpha}}}.$$

Thus, we can safely conclude that an increase in A reduces U.

To the extent that the transitionary dynamics in response to a change in either λ, A, or J are fast, business cycle fluctuations in unemployment and vacancies can be studied by performing the steady-state analysis described in the previous paragraph. Assuming that this is the case, we can make the following tentative hypothesis. First, since there is a strong negative empirical relationship between vacancies and unemployment between 2001:1 and 2007:12, fluctuations in the value of a filled job J, together with a relatively constant separation rate λ and a relatively constant matching productivity A, are the most likely scenario for explaining this period. Second, significant changes in A and/or λ are necessary for explaining the substantial deviations from the Beveridge curve observed after 2007:12, especially after 2009:6 (see figure 1). A key issue will be to determine the behavior of the matching efficiency parameter A during this later period. Another key issue will be to evaluate the contribution of changes in A, λ, and J to the unemployment and vacancy dynamics observed during this later period. Addressing these issues will be the focus of the next two subsections.

Before proceeding, it will be convenient to rewrite equations 7–8 as follows:

$$9)\quad U_t = \frac{\lambda_t}{\lambda_t + A_t\left(\frac{V_t}{U_t}\right)^{1-\alpha}},$$

$$10)\quad k = A_t\left(\frac{U_t}{V_t}\right)^{\alpha} J_t.$$

This makes explicit the assumption that the economy at any month t can be safely described by the steady-state equations 7–8, an assumption that will be maintained throughout the rest of this section.

Constant separation rate

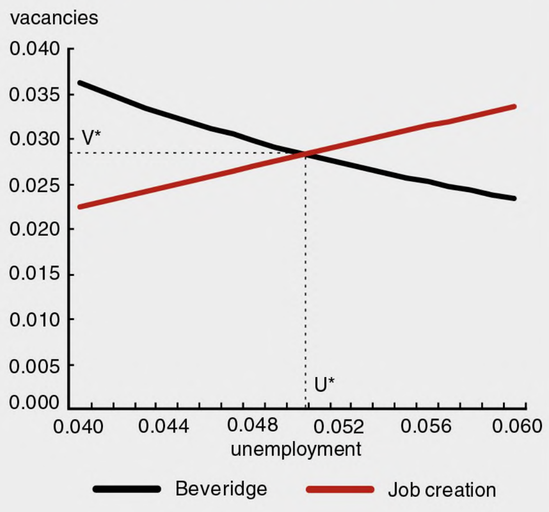
Shimer (2005) has argued that the separation rate λ does not play an important role in generating unemployment fluctuations. For this reason, I follow Barlevy (2011) and consider in this section that the separation rate λ is constant over time. Under this assumption, I use the model described by equations 9–10 to measure the time paths for the efficiency parameter A_t and the value of a job J_t.

In what follows, I set the separation rate λ to 0.042, which is equal to the average employment-to-unemployment transition rate plus the average employment-to-nonparticipation transition rate between 2001:1 and 2007:12. From equation 9, we have, for any two months i and j, that

$$11)\quad 1-\alpha_{j,i} = \left\{\ln\left[\frac{\lambda}{U_j} - \lambda\right] - \ln\left[\frac{\lambda}{U_i} - \lambda\right]\right\} \Big/ \left\{\ln\left(\frac{U_i}{V_i}\right) - \ln\left(\frac{U_j}{V_j}\right)\right\}.$$

FIGURE 2

Beveridge and job creation curves



Within the period 2001:1–2007:12 (which is a period with relatively constant matching productivity A), we can thus select the month i with the largest U/V ratio and the month j with the smallest U/V ratio and use them to get an estimate for α from equation 11.[4] These months happen to be i = 2003:6 and j = 2001:1. The estimated value of α turns out to be 0.4915.[5]

Equation 9 can also be used to measure the matching efficiency at month t as follows:

$$12)\quad A_t = \left[\frac{\lambda}{U_t} - \lambda\right]\left(\frac{U_t}{V_t}\right)^{1-\alpha}.$$

Using the above value of α and averaging the values of A_t between 2001:1 and 2007:12 obtained from equation 12 gives an estimate of A = 1.06.

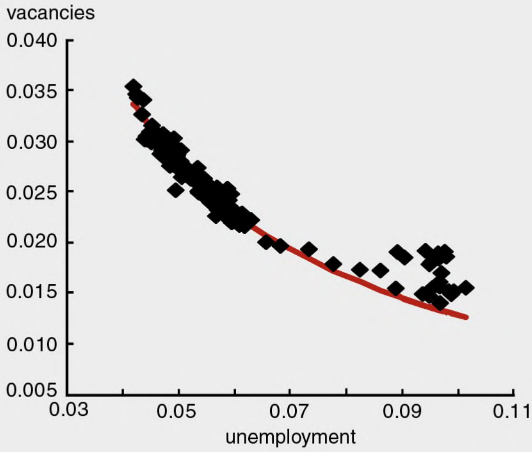
Using this constant value for A, we can then use equation 9 to construct the vacancies predicted by the model economy (conditional on the observed unemployment rate) as follows:

$$13)\quad V_t = \left\{\left[\frac{\lambda}{U_t} - \lambda\right]\frac{1}{A}\right\}^{\frac{1}{1-\alpha}} U_t.$$

The predicted vacancies are shown by the red line in figure 3. We see that under a constant matching efficiency parameter A, the model does a good job at reproducing the behavior of vacancies between 2001:1 and 2007:12. However, beginning in 2009, the model

FIGURE 3

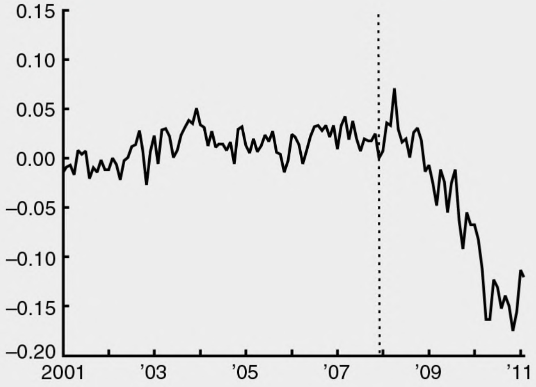
Predicted Beveridge curve



Note: Vacancies and unemployment are normalized by total labor force.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 4

Matching efficiency (2 states, constant separation rate)



Note: The log of A_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

fails to keep track of the data using a constant A. This suggests that the matching efficiency parameter A may have experienced substantial changes in this later period. To show that this could be the case, figure 4 reports the values for A_t (in logs) measured from equation 12 for the whole sample period (the vertical line corresponds to 2007:12, that is, to the start of the past recession). We see that the matching efficiency was relatively stable before 2008:1. However, starting in 2008:1, the matching efficiency has fluctuated quite substantially. In particular, we see that after an initial increase, the matching efficiency has been decreasing continuously, reaching a cumulative drop of 17.5 percent by 2010:11.

Normalizing the cost of posting a vacancy k to one and using the path for A_t already found, equation 10 can be used to construct a time series for the value of a job J_t. In particular, we have that

$$14)\ J_t = \frac{k}{A_t}\left(\frac{V_t}{U_t}\right)^{\alpha}.$$

Figure 5 reports the evolution of the value of a job between 2001:1 and 2011:2 (in logs). We see that J_t dropped quite substantially during the recession: Between 2007:12 and 2009:8, the value of a job declined by 68 percent.[6]

I now turn to evaluate the contributions of changes in A_t and J_t to the dynamics of unemployment and vacancies since the beginning of the recession. In order to do this, I compute adjusted unemployment U_t^* and adjusted vacancies V_t^* using equations 9–10 under the assumption that $A_t = A_{2007:12}$, for every month t. That is, I let the value of a job J_t evolve as in figure 5 (that is, as in the data) but fix the matching productivity to the value that it had at the start of the recession. For this reason, U_t^* and V_t^* measure the unemployment rates and vacancies that would have been obtained had the matching productivity remained constant at its December 2007 level but the path for the value of a job J_t had remained the same. Observe that in a full-blown model (in which J_t is endogenously determined), a change in the path for A_t would generally affect the path for J_t. As a consequence, comparing (U_t^*, V_t^*) with (U_t, V_t) cannot be strictly interpreted as describing the total effects of variations in A_t; it should be interpreted as describing the *conditional* effects of A_t (that is, conditional on the observed path for J_t). In a full-blown model, the variations in A_t would have to be accompanied by variations in other variables (for example, in the bargaining power of workers) in order to obtain an unchanged path for J_t.[7]

Figure 6 shows the path for U_t^* (labeled "constant A") and for U_t (labeled "variable A"). We see that through 2009:1, the productivity of the matching function did not play an important role in the unemployment dynamics observed (both paths are quite similar). However, starting in mid-2009, we see that the decline in matching productivity reported in figure 4 played a significant role in generating a significantly larger unemployment

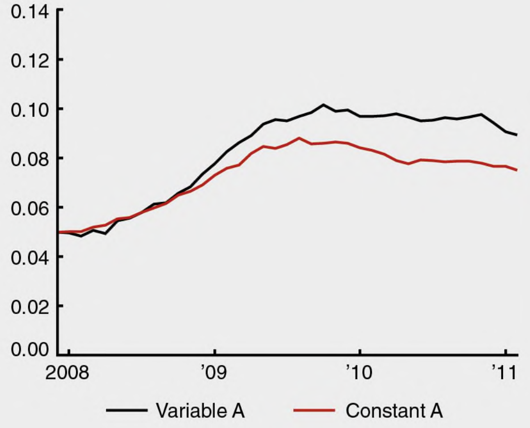
FIGURE 5

Value of a job (2 states, constant separation rate)

Note: The log of J_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 6

Effects on unemployment rate (2 states, constant separation rate)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

rate. In particular, we see that by February 2011 the unemployment rate would have been 7.5 percent instead of 8.9 percent had the matching productivity remained constant at its beginning-of-recession level.

Figure 7 reports the paths for V_t^* (labeled "constant A") and for V_t (labeled "variable A"). We also see that through 2009:1, changes in the productivity of the matching function had negligible effects on vacancies. However, by mid-2009 both paths start to diverge, and we see that by February 2011 vacancies would have been 1.6 percent instead of 1.9 percent had the matching productivity remained constant at its 2007:12 level.

Variable separation rate

In this section, I allow the separation rate to vary over time. Figure 8 reports the separation rate between 2001:1 and 2011:2 that is obtained from the Bureau of Labor Statistics' *Current Population Survey* (CPS) data (once again, the vertical line depicts the beginning of the last recession). We see that early on in the recession the separation rate increased quite significantly, reaching 4.9 percent by 2009:1, but that it subsequently trended down toward its pre-recession level.

Given the data on U_t and V_t and the separation rate λ_t reported in figure 8, I compute the matching efficiency A_t from the following equation:

$$15)\ A_t = \left[\frac{\lambda_t}{U_t} - \lambda_t\right]\left(\frac{U_t}{V_t}\right)^{1-\alpha},$$

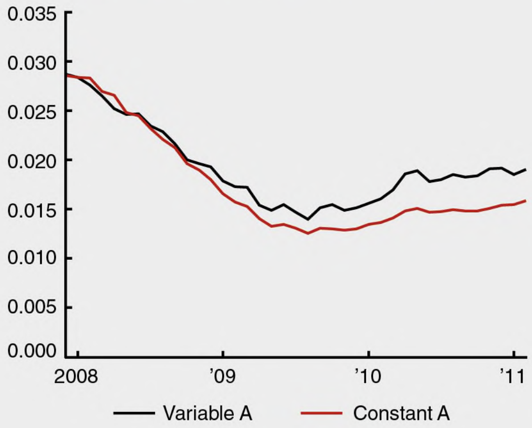
which is analogous to equation 12, except that λ_t is allowed to vary over time. The resulting path for the matching productivity A_t is reported in figure 9. We see that contrary to figure 4, we now observe large fluctuations in A_t previous to the start of the recession. Another difference is that there is a sharp increase in matching productivity early on in the recession that compensates for the 2009:1 spike in the separation rate. Also, we see that starting in 2009:2, the matching productivity trends down much more sharply than in figure 4.

The value of filled jobs J_t is computed from equation 14 using the A_t values obtained from equation 15. The resulting path is reported in figure 10. We see that this path is not very different from that in figure 5.

Figures 11 and 12 explore the conditional contribution to unemployment and vacancies dynamics of the matching productivity A_t, the separation rate λ_t, and the value of a job J_t. In particular, I compute adjusted unemployment U_t^* and adjusted vacancies V_t^* using equations 9–10 under the assumption that $A_t = A_{2007:12}$ and $\lambda_t = \lambda_{2007:12}$, for every month t. That is, I let the value of a job J_t evolve as in figure 10 (that is, as in the data), but I fix the matching productivity to the value that it had at the beginning of the recession $A_{2007:12}$ and fix the separation rate to the value that it had at the beginning of the recession $\lambda_{2007:12}$. In other words, U_t^* and V_t^* measure the unemployment rates and vacancies that would have been obtained had the matching productivity and the separation rate remained constant at their December 2007 levels.

FIGURE 7

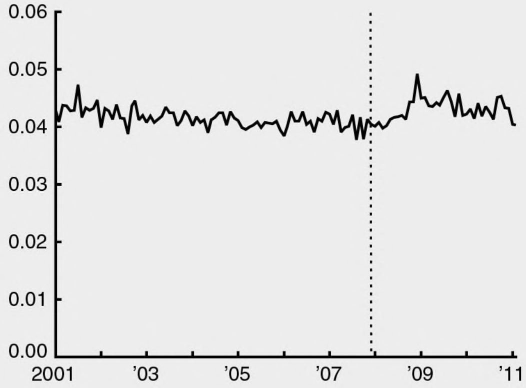
Effects on vacancies (2 states, constant separation rate)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 8

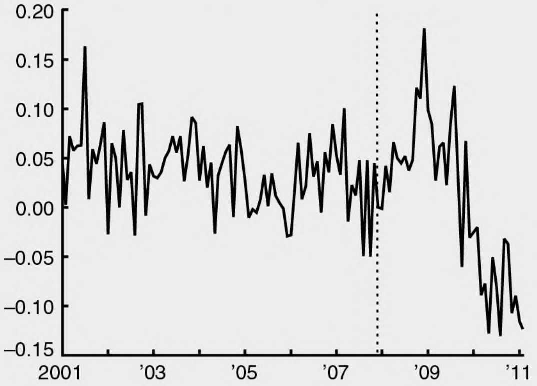
Separation rate



Note: The separation rate is equal to the sum of the employment-to-unemployment and employment-to-nonparticipation transition rates.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey*.

FIGURE 9

Matching efficiency (2 states, variable separation rate)



Note: The log of A_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 10

Value of a job (2 states, variable separation rate)

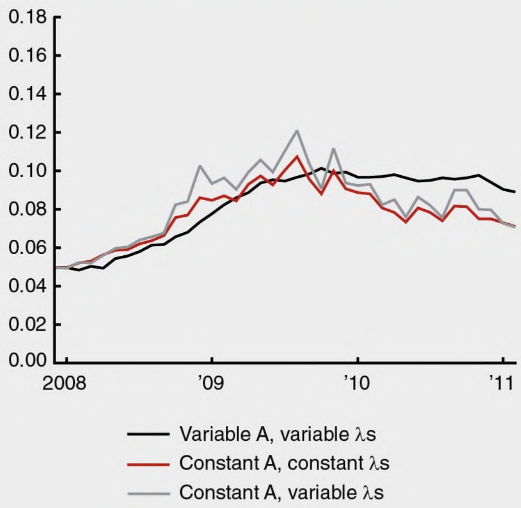
Note: The log of J_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

I also compute adjusted unemployment U_t^{**} and adjusted vacancies V_t^{**} using equations 9–10 under the assumption that $A_t = A_{2007:12}$, for every month t (but I let the separation rate λ_t vary as in the data). That is, I let the value of a job J_t evolve as in figure 10 and the separation rate λ_t evolve as in figure 8 but fix the matching productivity to the value that it had at the beginning of the recession $A_{2007:12}$. In other words, U_t^{**} and V_t^{**} measure the unemployment rates and vacancies that would have been obtained had the matching productivity remained constant at its December 2007 level.

In figure 11, U_t^* is labeled "constant A, constant λ," U_t^{**} is labeled "constant A, variable λ," and U_t

FIGURE 11

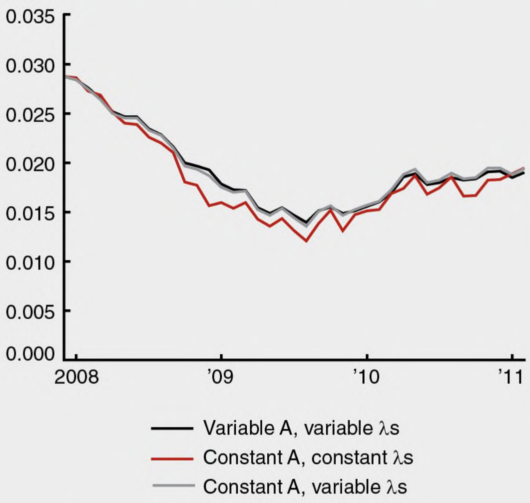
Effects on unemployment rate (2 states, variable separation rate)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 12

Effects on vacancies (2 states, variable separation rate)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

is labeled "variable A, variable λ." We see that U_t^* increases in the early part of the period and decreases during the second part, mirroring the evolution of the value of a job J_t described in figure 10. In turn, U_t^{**} increases much more than U_t^* early on in the recession because of the early increase in the separation rate λ_t depicted in figure 8, but as the separation rate reverts toward its beginning-of-recession level, U_t^{**} starts to behave very much like U_t^*. Finally, since the difference between U_t^{**} and U_t is solely due to changes in the matching productivity, we see that the large increase in matching productivity early on in the recession (reported in figure 9) played an important role in keeping unemployment relatively low. However, the large drop in matching productivity since early 2009 significantly contributed to maintaining an unemployment rate of more than 9 percent.

In turn, figure 12 shows that the matching productivity doesn't play a crucial role in vacancy dynamics. However, the large increase in the separation rate in the early part of the recession played a noticeable role in keeping vacancies relatively high early on in the recession.

The case of three labor market states

The model used in the previous section had two labor market states: employment and unemployment. In this section, I allow workers to be in a third labor market state: nonparticipation (that is, out of the labor force). A main reason for doing this is that in the CPS data between 2001:1 and 2007:12, the total number of people transitioning from nonparticipation to employment is almost twice as large as the total number of people transitioning from unemployment to employment (see figure 13), although the differences have become much smaller since the start of the past recession. By considering only two market states, the analysis in the previous section completely missed these transitions. Another reason for introducing three labor market states into the model is that with two labor market states, it is not clear what separation rates to consider: separations into unemployment or separations into both unemployment and nonparticipation? Explicitly introducing three labor markets states avoids this type of issue. More generally, introducing three labor market states allows me to address worker flows data in a more satisfactory way.

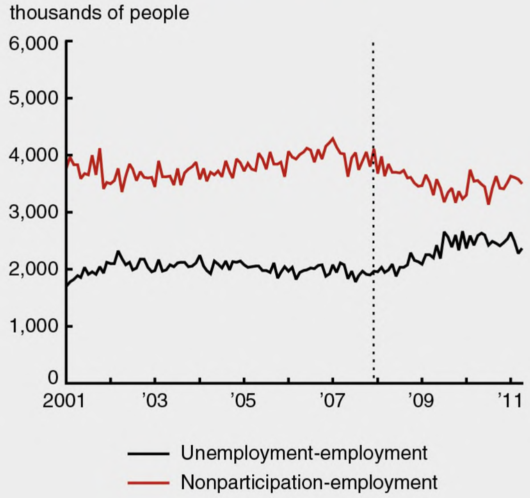
In this section, I follow Barnichon and Figura (2010) and assume that the matching function solely describes transitions from unemployment into employment. In particular, I assume that the matching function is given by

$$16)\quad M_t = A_t U_t^{\alpha} V_t^{1-\alpha},$$

where M_t are the total flows from unemployment into employment, U_t is unemployment, V_t are vacancies, and $0 < \alpha < 1$.

FIGURE 13

Flows into employment



Source: Bureau of Labor Statistics' *Current Population Survey*.

The transition rate from employment to unemployment λ_t^{EU}, the transition rate from employment to nonparticipation λ_t^{EN}, the transition rate from unemployment to nonparticipation λ_t^{UN}, the transition rate from nonparticipation to unemployment λ_t^{NU}, and the transition rate from nonparticipation to employment λ_t^{NE} are assumed to be exogenous to the model.

The evolution of workers across labor market states is then given by the following equations:

$$17)\ E_{t+1} = E_t + A_t U_t^{\alpha} V_t^{1-\alpha} + \lambda_t^{NE} N_t - (\lambda_t^{EU} + \lambda_t^{EN}) E_t,$$

$$18)\ U_{t+1} = U_t + \lambda_t^{EU} E_t + \lambda_t^{NU} N_t - \lambda_t^{UN} U_t^{\alpha} - A_t U_t^{\alpha} V_t^{1-\alpha},$$

$$19)\ N_{t+1} = N_t + \lambda_t^{EN} E_t + \lambda_t^{UN} U_t - (\lambda_t^{NE} + \lambda_t^{NU}) N_t.$$

Equation 17 states that next-period employment is equal to current employment, plus all transitions into employment (either from unemployment or nonparticipation), minus total separations (either to unemployment or nonparticipation). Equation 18 states that next-period unemployment is equal to current unemployment, plus all transitions into unemployment (either from employment or nonparticipation), minus all transitions out of unemployment (either to employment or nonparticipation). Equation 19 states that next-period nonparticipation is equal to current nonparticipation, plus all transitions into nonparticipation (either from employment or unemployment), minus all transitions out of nonparticipation (either to employment or unemployment).

In what follows, total population will be normalized to one, that is,

$$20)\ E_t + U_t + N_t = 1,$$

for every period t.

Similar to the previous section, the following free-entry condition must be satisfied:

$$21)\ k = \left(\frac{A_t U_t^{\alpha} V_t^{1-\alpha} + \lambda_t^{NE} N_t}{V_t} \right) J_t.$$

Observe that from the point of view of a firm, the probability of filling a vacancy is equal to $(M_t + \lambda_t^{NE} N_t)/V_t$, because matches can be formed with workers either coming from unemployment or from nonparticipation.

Given the total number of workers unemployed at date zero U_0 and the total number of workers that are nonparticipants at date zero N_0, the model generates an endogenous path for $\{M_t, V_t, U_{t+1}, N_{t+1}\}_{t=0}^{\infty}$.

Steady state

Assuming a constant matching productivity A, a constant value of a job J, and constant transition rates λ, a steady state of the model economy can be defined as an initial unemployment level $U_0 = U$ and an initial nonparticipation level N_0, such that the endogenous path for $\{M_t, V_t, U_{t+1}, N_{t+1}\}_{t=0}^{\infty}$ that the model generates is constant over time.

From equations 17–19, we have that the conditions that a steady state (U, N, V) must satisfy are the following:

$$22)\ (\lambda^{EU} + \lambda^{EN})(1 - U - N) = AU^{\alpha}V^{1-\alpha} + \lambda^{NE} N,$$

$$23)\ (\lambda^{NE} + \lambda^{NU} + \lambda^{EN})\ N = \lambda^{EN}(1 - U) + \lambda^{UN} U,$$

$$24)\ k = \left(\frac{AU^{\alpha}V^{1-\alpha} + \lambda^{NE} N}{V} \right) J.$$

Similar to the previous section, it will be convenient to rewrite these equations as:

$$25)\ (\lambda_t^{EU} + \lambda_t^{EN})(1 - U_t - N_t) = A_t U_t^{\alpha} V_t^{1-\alpha} + \lambda_t^{NE} N_t,$$

$$26)\quad N_t = \frac{\lambda_t^{EN}(1-U_t) + \lambda_t^{UN}U_t}{\left(\lambda_t^{NE} + \lambda_t^{NU} + \lambda_t^{EN}\right)},$$

$$27)\quad \frac{kV_t}{J_t} = A_t U_t^{\alpha} V_t^{1-\alpha} + \lambda_t^{NE} N_t.$$

This makes explicit the assumption that the economy at any month t can be safely described by the steady-state equations 22–24, an assumption that will be maintained throughout the following two subsections.

Constant transition rates

Figure 14 shows the transition rates λ_t^{UN}, λ_t^{EU}, λ_t^{EN} λ_t^{NE}, and λ_t^{NU}, in logs, normalized by their average value for the period 2001:1–2007:12. We see that these transition rates were relatively stable prior to 2007:12. However, we see that with the onset of the recession, there was a significant drop in the transition rate from nonparticipation to employment λ_t^{NE}, a drop in the transition rate from unemployment to nonparticipation λ_t^{UN}, a large increase in the transition rate from nonparticipation to unemployment λ_t^{NU}, and a large increase in the transition rate from employment to unemployment λ_t^{EU}. In turn, the transition rate from employment to nonparticipation was not significantly affected.

Based on figure 14, and similar to the previous section, here I assume that the transition rates λ_t^{UN}, λ_t^{EU}, λ_t^{EN} λ_t^{NE}, and λ_t^{NU} are constant over the period 2001:1–2007:12. Taking simple averages over this period gives the following values:

$$28)\quad \lambda^{UN} = 0.2258,$$

$$29)\quad \lambda^{EU} = 0.0132,$$

$$30)\quad \lambda^{EN} = 0.0281,$$

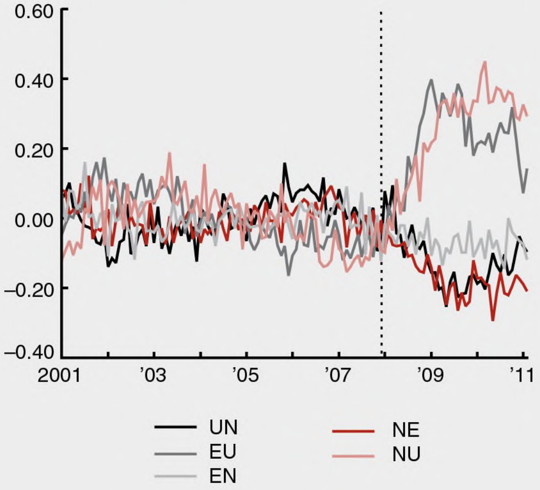
$$31)\quad \lambda^{NE} = 0.0505,$$

$$32)\quad \lambda^{NU} = 0.0253.$$

Substituting equation 26 in equation 25 under the assumption of constant λ values, we get that

FIGURE 14

Transition rates (λs)



Note: All transition rates are reported in logs and normalized to zero at the start of the last recession.
Source: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey*.

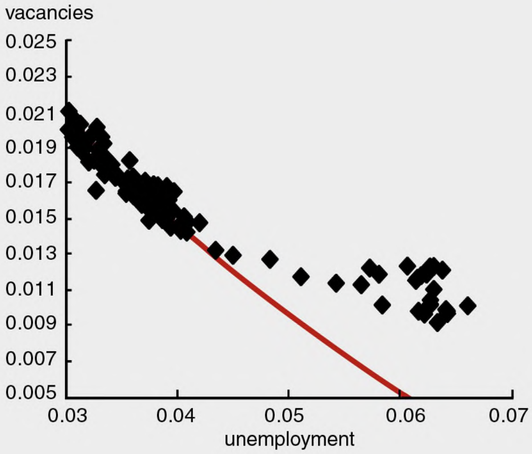
$$33)\quad A_t = \frac{1}{U_t^{\alpha}V_t^{1-\alpha}} \left\{ \begin{array}{l} \left(\lambda^{EU} + \lambda^{EN}\right)\left(1 - U_t - \left[\dfrac{\lambda^{EN}(1-U_t) + \lambda^{UN}U_t}{\left(\lambda^{NE} + \lambda^{NU} + \lambda^{EN}\right)}\right]\right) \\ -\lambda^{NE}\left[\dfrac{\lambda^{EN}(1-U_t) + \lambda^{UN}U_t}{\left(\lambda^{NE} + \lambda^{NU} + \lambda^{EN}\right)}\right] \end{array} \right\}.$$

Assuming that A is constant over the period 2001:1–2007:12, we can then use two months i and j within this period to get an estimate of α as follows:

$$34)\quad 1-\alpha_{ij} = \left[\frac{1}{\ln\left(\frac{V_i}{U_i}\right) - \ln\left(\frac{V_j}{U_j}\right)}\right] \times \{\ln \left[\begin{array}{l} \left(\lambda^{EU} + \lambda^{EN}\right) \\ \left(1 - U_i - \left[\dfrac{\lambda^{EN}(1-U_i) + \lambda^{UN}U_i}{\left(\lambda^{NE} + \lambda^{NU} + \lambda^{EN}\right)}\right]\right) \\ -\lambda^{NE}\left[\dfrac{\lambda^{EN}(1-U_i) + \lambda^{UN}U_i}{\left(\lambda^{NE} + \lambda^{NU} + \lambda^{EN}\right)}\right] \end{array} \right]$$

FIGURE 15

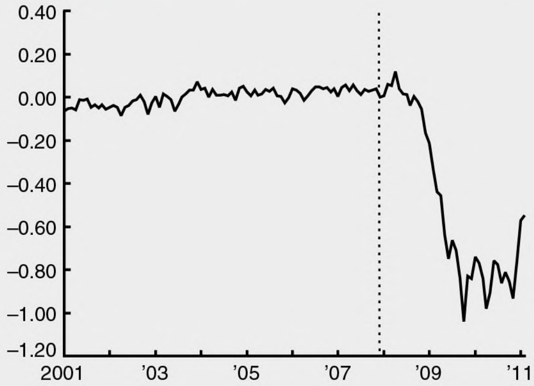
Predicted Beveridge curve



Note: Vacancies and unemployment are normalized by total civilian noninstitutional population (16 years and older).
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 16

Matching efficiency (3 states, steady state, constant transition rates)



Note: The log of A_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

$$-\ln\left[\begin{array}{l}(\lambda^{EU}+\lambda^{EN})\left(1-U_j-\left[\frac{\lambda^{EN}(1-U_j)+\lambda^{UN}U_j}{(\lambda^{NE}+\lambda^{NU}+\lambda^{EN})}\right]\right)\\ -\lambda^{NE}\left[\frac{\lambda^{EN}(1-U_j)+\lambda^{UN}U_j}{(\lambda^{NE}+\lambda^{NU}+\lambda^{EN})}\right]\end{array}\right]$$

$$+\ln(U_j)-\ln(U_i)\}.$$

Picking $i = 2003{:}6$ and $j = 2001{:}1$, which are the months with the largest and smallest V/U ratio, respectively, gives an estimate of $\alpha = 0.16$.[8]

In turn, equation 33 can be used to measure the matching efficiency A_t at month t. Using the above value of α and averaging the values of A_t between 2001:1 and 2007:12 obtained from equation 33 gives an estimate of $A = 0.4533$.

Using this constant value for A, I can then use equation 33 to construct the vacancies predicted by the model economy (conditional on observed unemployment) as follows:

$$35)\quad V_t=\left(\frac{1}{AU_t^{\alpha}}\right)\times\left\{\begin{array}{l}(\lambda^{EU}+\lambda^{EN})\\ \left(1-U_t-\left[\frac{\lambda^{EN}(1-U_t)+\lambda^{UN}U_t}{(\lambda^{NE}+\lambda^{NU}+\lambda^{EN})}\right]\right)\\ -\lambda^{NE}\left[\frac{\lambda^{EN}(1-U_t)+\lambda^{UN}U_t}{(\lambda^{NE}+\lambda^{NU}+\lambda^{EN})}\right]\end{array}\right\}^{\frac{1}{1-\alpha}}.$$

Figure 15 reports unemployment as a fraction of total population and vacancies as a fraction of total population between 2001:1 and 2011:2 (black dots), as well as the vacancies predicted by equation 35. We see that the steady state of the model with three labor market states, constant transition rates, and a constant A provides a good fit to the data through 2007:12. However, since the start of the latest recession there have been large deviations from the stable Beveridge curve predicted by the model. This indicates that the matching efficiency parameter A_t must have experienced significant changes since then. Figure 16 shows that this has been the case. It reports the matching efficiency levels obtained by equation 33 between 2001:1 and 2011:2. We see that before 2007:12, the matching efficiency had been fairly stable, but it plummeted with the onset

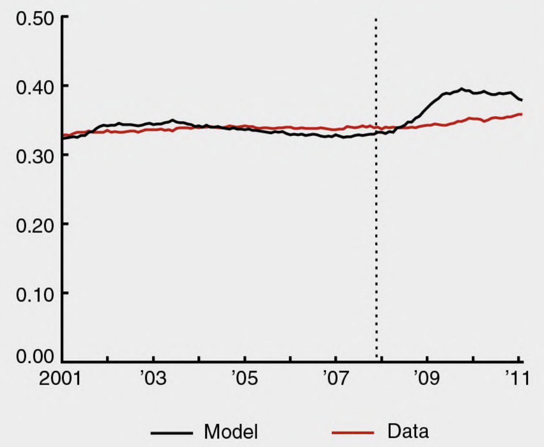
FIGURE 17

Value of a job
(3 states, steady state, constant transition rates)

Note: The log of J_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 18

Nonparticipation
(3 states, steady state, constant transition rates)



Note: Nonparticipation is normalized by total civilian noninstitutional population (16 years and older).
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

of the recession. Observe that the magnitude of the fall is much larger than in figure 4 (p. 151).

Substituting equation 26 in equation 27 we get that

$$36)\ J_t = kV_t\left(\frac{A_t U_t^{\alpha} V_t^{1-\alpha} + \lambda^{NE}}{\left[\frac{\lambda^{EN}(1-U_t)+\lambda^{UN}U_t}{(\lambda^{NE}+\lambda^{NU}+\lambda^{EN})}\right]}\right)^{-1}.$$

Figure 17 reports the path for J_t thus measured. We see that it is very similar to that in figure 5 (p. 152), indicating that having three labor market states does not significantly affect the measurement of the value of a job.

Before decomposing the effects of the matching efficiency parameter A_t and the value of a job J_t, I would like to point out that the results that follow should be taken with a grain of salt. While I selected the paths for A_t and J_t to reproduce the observed path for U_t and V_t (given the restrictions imposed by equations 33 and 36), I made no attempt to reproduce the path for nonparticipation N_t, which according to the model is given by

$$37)\ N_t = \frac{\lambda^{EN}(1-U_t)+\lambda^{UN}U_t}{(\lambda^{NE}+\lambda^{NU}+\lambda^{EN})}.$$

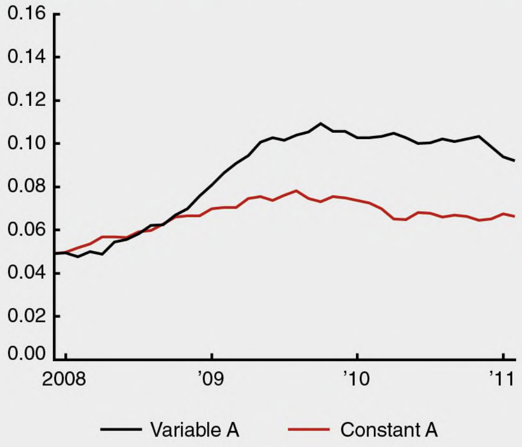
Figure 18 reports the path for nonparticipation in U.S. data and the path for N_t given by equation 37. We see that the model does a reasonable job at reproducing the path for N_t before 2007:12, but that it largely overpredicts nonparticipation after that. This suggests that either the assumption of constant transition rates or the assumption that the economy is always at the steady state of the model fails. I return to this issue in the next section.

Similar to the previous section, I compute adjusted unemployment U_t^*, adjusted nonparticipation N_t^*, and adjusted vacancies V_t^* from equations 25–27 under the assumption that $A_t = A_{2007:12}$, for every month t. That is, I let the value of a job J_t evolve as in figure 17 (that is, as in the data) but fix the matching productivity to the value that it had at the start of the recession. In other words, U_t^*, N_t^*, and V_t^* measure the unemployment, nonparticipation, and vacancies that would have been obtained had the matching productivity remained constant at its December 2007 level but the value of a job had evolved as observed.

The version of the model with constant transition rates delivers the following results.[9] Similar to figure 6 (p. 152), figure 19 indicates that starting in mid-2009, the decline in matching productivity reported in figure 16 played an important role in generating a large unemployment rate. This version of the model also indicates that by February 2011, the unemployment rate would have been 6.4 percent instead of 8.9 percent had matching productivity remained constant at its beginning-of-recession level. Figure 20 shows that

FIGURE 19

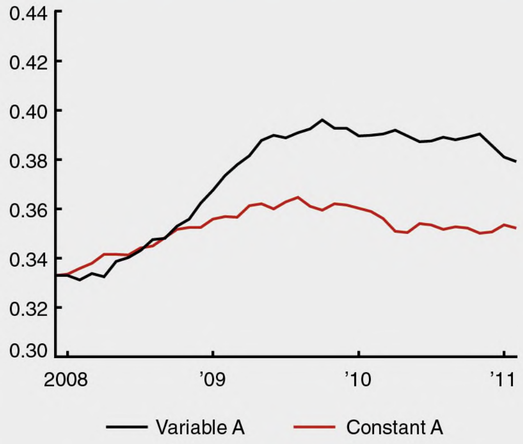
Effects on unemployment rate (3 states, steady state, constant transition rates)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 20

Effects on nonparticipation (3 states, steady state, constant transition rates)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

the effects of matching productivity on nonparticipation are very similar to those on unemployment. Finally, similar to figure 7 (p. 153), figure 21 shows that the effects of matching efficiency on vacancies are negligible.

Variable transition rates

In this section, instead of assuming that transition rates are constant, I allow them to fluctuate as in figure 14. Given data on U_t and all transition rates, I compute matching efficiency as follows:

$$38)\ A_t = \frac{1}{U_t^{\alpha} V_t^{1-\alpha}} \times \left\{ \begin{array}{l} \left(\lambda_t^{EU} + \lambda_t^{EN}\right) \\ \left(1 - U_t - \left[\frac{\lambda_t^{EN}\left(1-U_t\right) + \lambda_t^{UN} U_t}{\left(\lambda_t^{NE} + \lambda_t^{NU} + \lambda_t^{EN}\right)}\right]\right) \\ -\lambda_t^{NE}\left[\frac{\lambda_t^{EN}\left(1-U_t\right) + \lambda_t^{UN} U_t}{\left(\lambda_t^{NE} + \lambda_t^{NU} + \lambda_t^{EN}\right)}\right] \end{array} \right\},$$

an expression obtained from substituting equation 26 in equation 25. Figure 22 reports the path for A_t thus obtained. We observe huge differences from figure 16. Instead of relatively stable behavior before 2007:12 followed by a large drop, we observe significant volatility throughout the sample period and a large increase after 2007:12. These differences indicate that the predictions of the model rely critically on whether transition rates are assumed to be constant or not.

The value of a job J_t is measured as

$$39)\ J_t = kV_t \left(A_t U_t^{\alpha} V_t^{1-\alpha} + \lambda_t^{NE} \left[\frac{\lambda_t^{EN}\left(1-U_t\right) + \lambda_t^{UN} U_t}{\left(\lambda_t^{NE} + \lambda_t^{NU} + \lambda_t^{EN}\right)} \right] \right)^{-1}$$

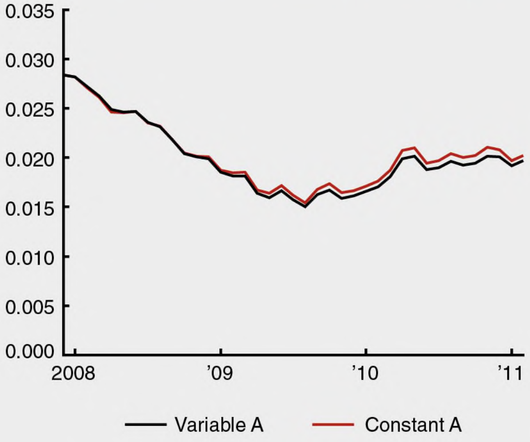
and reported in figure 23. We see that the qualitative behavior is similar to figure 17; however, the drop in J_t after the start of the past recession is now somewhat larger.

Before turning to the decomposition of the different effects, I revisit the issue of how well the model is able to reproduce the path for nonparticipation, a path that has not been targeted in the calibration. Figure 24 reports the path for nonparticipation in U.S. data and the path for N_t from equation 26. We see that contrary to figure 18, the model now does a reasonable job at reproducing the path for N_t throughout the sample period. In principle, this should be a reason for having more confidence in the results obtained in this section.

In order to decompose the different effects, I compute adjusted unemployment U_t^*, adjusted nonparticipation N_t^*, and adjusted vacancies V_t^* from equations 25–27 under the assumption that $A_t = A_{2007:12}$, $\lambda_t^{UN} = \lambda_{2007:12}^{UN}$, $\lambda_t^{EU} = \lambda_{2007:12}^{EU}$, $\lambda_t^{EN} = \lambda_{2007:12}^{EN}$, $\lambda_t^{NE} = \lambda_{2007:12}^{NE}$, and $\lambda_t^{NU} = \lambda_{2007:12}^{NU}$, for every month t. That is, I let the

FIGURE 21

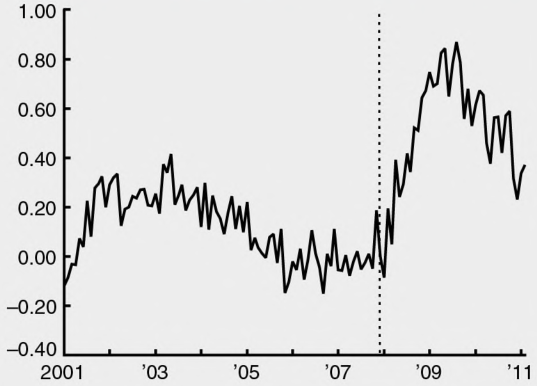
Effects on vacancies (3 states, steady state, constant transition rates)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 22

Matching efficiency (3 states, steady state, variable transition rates)



Note: The log of A_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

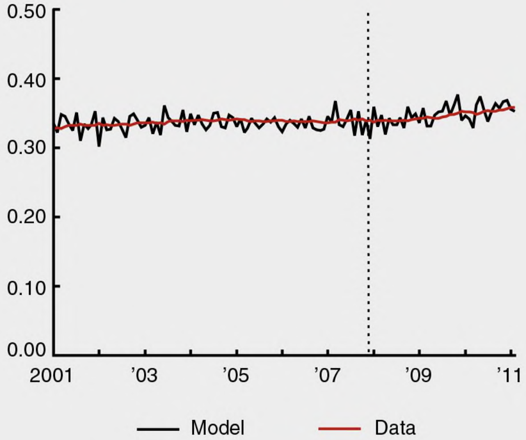
FIGURE 23

Value of a job (3 states, steady state, variable transition rates)

Note: The log of J_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 24

Nonparticipation (steady state, variable transition rates)

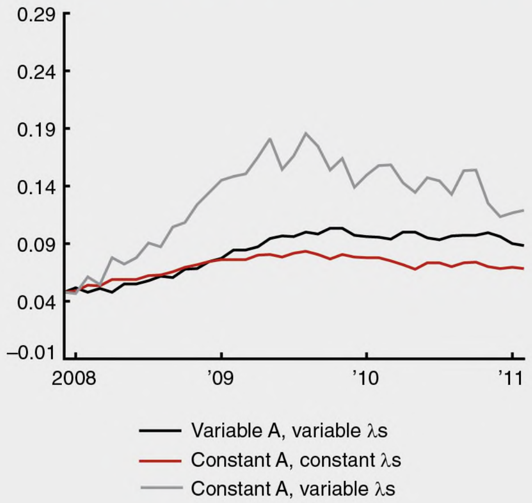


Note: Nonparticipation is normalized by total civilian noninstitutional population (16 years and older).
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

value of a job J_t evolve as in figure 23 (that is, as in the data) but fix the matching productivity and the transition rates to the values that they had at the beginning of the recession. In other words, U_t^*, N_t^*, and V_t^* measure the unemployment, nonparticipation, and vacancies that would have been obtained had the matching productivity and transition rates remained constant at their December 2007 levels but J_t had evolved as it did.

FIGURE 25

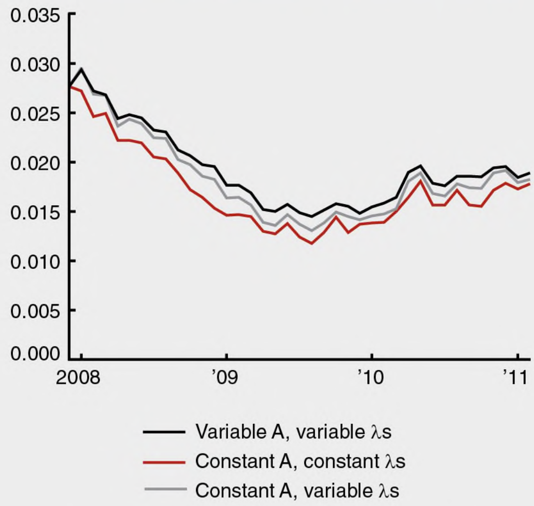
Effects on unemployment rate (3 states, steady state, variable transition rates)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 26

Effects on vacancies (3 states, steady state, variable transition rates)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

Also, I compute adjusted unemployment U_t^{**}, adjusted nonparticipation N_t^{**}, and adjusted vacancies V_t^{**} from equations 25–27 under the assumption that $A_t = A_{2007:12}$, for every month t (but letting all λ_ts take their actual values). That is, I let the value of a job J_t evolve as in figure 23 and the transition rates evolve as in figure 14 (p. 156), but I fix the matching productivity to the value that it had at the beginning of the recession $A_{2007:12}$. In other words, U_t^{**}, N_t^{**}, and V_t^{**} measure the unemployment, nonparticipation, and vacancies that would have been obtained had the matching productivity remained constant at its December 2007 level but all transition rates and J_t had evolved as they did.

Figure 25 shows $U_t^{*} / (E_t^{*} + U_t^{*})$ ("constant A, constant λ"), $U_t^{**} / (E_t^{**} + U_t^{**})$ ("constant A, variable λ"), and $U_t / (E_t + U_t)$ ("variable A, variable λ"). We see that despite the large drop in the value of a job J_t described in figure 23, $U_t^{*} / (E_t^{*} + U_t^{*})$ increased only moderately. In turn, $U_t^{**} (E_t^{**} + U_t^{**})$ increases by a huge amount, indicating that the large increases in the λ_t^{EU} and λ_t^{NU} observed after 2007:12 in figure 14 had a large negative impact in the labor market. Actually, the unemployment rate turned to increase only as described by $U_t / (E_t + U_t)$ because of the large increase in matching efficiency reported by figure 12 (p. 154).

Figure 26 shows that the increase in matching productivity A_t and the changes in transition rates played a noticeable role in keeping vacancies relatively high. In turn, figure 27 shows that the increase in matching productivity A_t played a crucial role in keeping nonparticipation (N) relatively low, since the changes in transition rates would have increased it quite substantially.

Transitionary dynamics

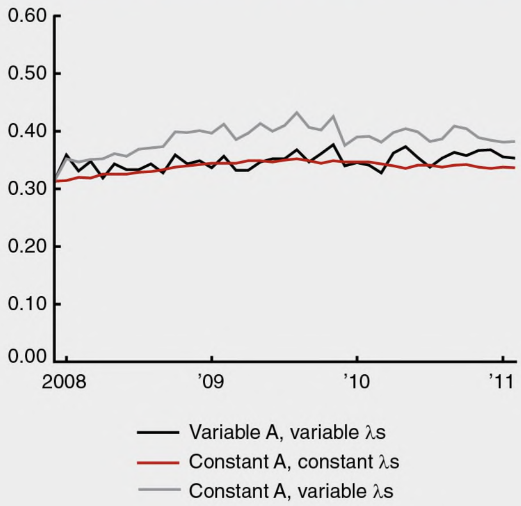
The previous section showed that introducing variable transition rates affects the results quite significantly and that it allows one to keep track of the behavior of nonparticipation much more closely. However, the analysis of the previous section suffered two drawbacks. First, while the calibration of the matching elasticity parameter α assumed constant matching efficiency and constant separation rates prior to 2007:12, we see from figures 14 (p. 156) and 22 that this is not quite the case. Second, the analysis assumed that the steady state of the model could be used to describe monthly data, while the large fluctuations in transition rates observed in figure 14 (p. 156) suggest that this may not be a good approximation. For these reasons, in this section I take a more direct approach to the calibration of the matching elasticity parameter α and perform the analysis without imposing that the model is always at its steady state. This allows me to evaluate to what extent this affects the results.

Observe from equation 16 that

$$40)\quad \ln\left(\frac{M_t}{V_t}\right) = \ln(A_t) + \alpha \ln\left(\frac{U_t}{V_t}\right).$$

FIGURE 27

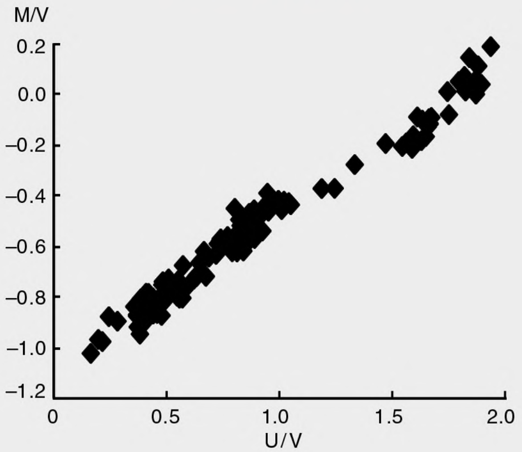
Effects on nonparticipation (3 states, steady state, variable transition rates)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 28

Matching function



Note: M/V is total matches per vacancy; U/V is the unemployment–vacancy ratio.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

In what follows, I identify M_t with the total number of workers that transition from unemployment to employment between months t and $t + 1$, as reported by the CPS. Figure 28 plots $\ln\left(\frac{M_t}{V_t}\right)$ against $\ln\left(\frac{U_t}{V_t}\right)$ for the whole sample period. We see a strong linear relation, suggesting that equation 40 provides a good description of the data with a relatively constant A_t. Fitting equation 40 using OLS (ordinary least squares) over the period 2001:1–2007:12 gives an estimate of $\alpha = 0.69$.

Given this estimated value of α, the path for the matching efficiency parameter A_t implied by equation 40 is reported in figure 29. We see that this path is completely different from that in figure 22. The matching productivity is much less variable and contrary to figure 22, displays a large drop after the start of the last recession, reaching by the end of the sample period a value 12 percent lower than in 2007:12. There is no doubt that measuring A_t directly from the matching function in equation 16 gives a very different picture from measuring it from the steady states of the model economy.

Figure 30 reports the value of a job J_t obtained from equation 27 using the matching efficiencies A_t obtained from equation 40 and reported in figure 29. The figure is very similar to figure 23, again indicating that the path for the value of a job is robust to the different ways of measuring it.

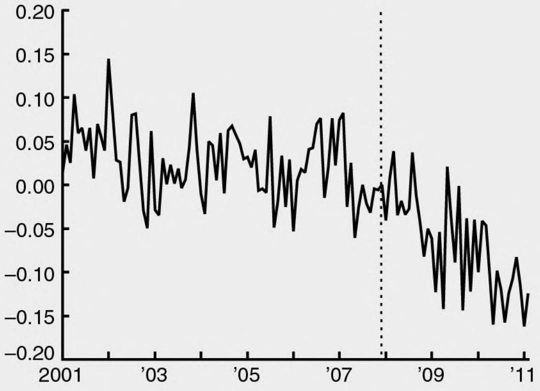
I now turn to evaluating the relative contributions of the value of filled jobs J_t, the matching efficiency parameter A_t, and the transition rates λ_t^{EU}, λ_t^{EN}, λ_t^{UN}, λ_t^{NU}, and λ_t^{NE} to unemployment dynamics since the beginning of the recession. For this purpose, I proceed as before and find a sequence E_t^*, U_t^*, N_t^*, and V_t^* that satisfies equations 17, 18, 19, and 21 under the assumption that $A_t = A_{2007:12}$, $\lambda_t^{UN} = \lambda_{2007:12}^{UN}$, $\lambda_t^{EU} = \lambda_{2007:12}^{EU}$, $\lambda_t^{EN} = \lambda_{2007:12}^{EN}$, $\lambda_t^{NE} = \lambda_{2007:12}^{NE}$, and $\lambda_t^{NU} = \lambda_{2007:12}^{NU}$, for every month t. That is, I let the value of a job J_t evolve as in figure 30 but fix the matching productivity and all transition rates to the values that they had in 2007:12 (that is, at the beginning of the recession). Similarly, as before, E_t^*, U_t^*, N_t^*, and V_t^*, describe the employment, unemployment, nonparticipation, and vacancies levels that would have obtained if the value of a job had been the only variable changing over time.

Also, I compute the $E_t^{**}, U_t^{**}, N_t^{**}$, and V_t^{**} that satisfy equations 17, 18, 19, and 21 under the assumption that $A_t = A_{2007:12}$, for every month t. That is, E_t^{**}, U_t^{**}, N_t^{**}, and V_t^{**} describe the employment, unemployment, nonparticipation, and vacancies levels that would have been obtained if the matching productivity parameter had remained constant at its December 2007 level, while all other variables (that is, J_t and all the λ values) had changed the way they did.

Figure 31 reports the paths for $U_t^* / (E_t^* + U_t^*)$ ("constant A, constant λ"), $U_t^{**} / (E_t^{**} + U_t^{**})$ ("constant A, variable λ"), and $U_t / (E_t + U_t)$ ("variable A, variable λ"). From the $U_t^* / (E_t^* + U_t^*)$ path, we see

FIGURE 29

Matching efficiency (3 states, transitional dynamics)



Note: The log of A_t is normalized to zero at the start of the past recession.

Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 30

Value of a job (3 states, transitional dynamics)

Note: The log of J_t is normalized to zero at the start of the past recession.

Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

that changes in the value of a job J_t played a very minor role in unemployment dynamics: The red line is roughly flat. From comparing the path for $U_t^{**}/(E_t^{**}+U_t^{**})$ with the path for $U_t^{*}/(E_t^{*}+U_t^{*})$, I conclude that changes in the transition rates λ played a crucial role in unemployment dynamics: The gray line is widely different from the red line. In fact, we see that changes in the transition rates λ accounted for most of the unemployment dynamics observed since the recession: The black line is very close to the gray line, indicating that the matching productivity played a minor role in the observed unemployment rate dynamics.

In turn, figure 32 reports the path for N_t^{*} ("constant A, constant λ"), N_t^{**} ("constant A, variable λ") and N_t ("variable A, variable λ"). From the N_t^{*} path, we see that far from accounting for the observed increase in nonparticipation, changes in the value of a job J_t would have accounted for a decrease in nonparticipation. The bulk of the increase in nonparticipation is accounted for by changes in transition rates, since changes in the matching productivity played a relatively minor role: The gray line is very close to the black line).

Finally, figure 33 shows that none of the changes in the matching efficiency parameter A_t or in the transition probabilities were important determinants of vacancies dynamics: The path for vacancies was mainly determined by J_t.

Nonparticipants compete for vacancies

This section describes and uses the most satisfactory specification of the model. Thus, it provides the main results of the paper. Observe that the model used in the previous section had three labor market states but only unemployed workers were inputs to the matching function: Nonparticipants made transitions to employment but without going through the matching function. I view this feature as a weakness of the previous specification of the model. The workers transitioning from nonparticipation to employment must be competing for the same vacancies as the workers transitioning from unemployment to employment and should therefore enter the matching function in a similar way. This section addresses this issue by modifying the matching function of the previous section accordingly. Introducing a more satisfactory specification for the matching function allows me to obtain better measurements of the matching efficiency.

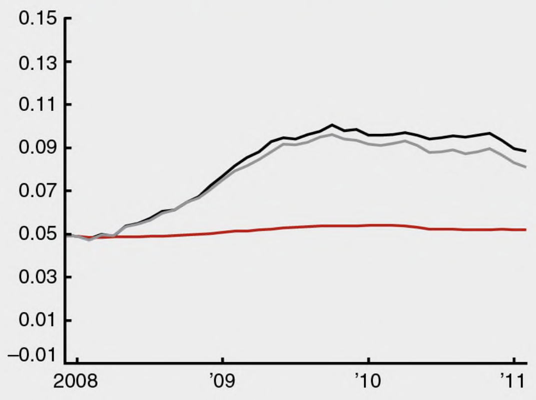
The matching function is now described as follows:

$$41)\quad M_t = A_t\left(U_t + \psi_t N_t\right)^{\alpha} V_t^{1-\alpha},$$

where M_t is the total number of matches, U_t is unemployment, N_t is nonparticipation, V_t is vacancies, A_t is the matching efficiency, $0 \le \psi_t \le 1$, and $0 < \alpha < 1$. Observe that ψ_t can be interpreted as the fraction of the total number of workers who report they are nonparticipants but search for jobs anyway. Alternatively,

FIGURE 31

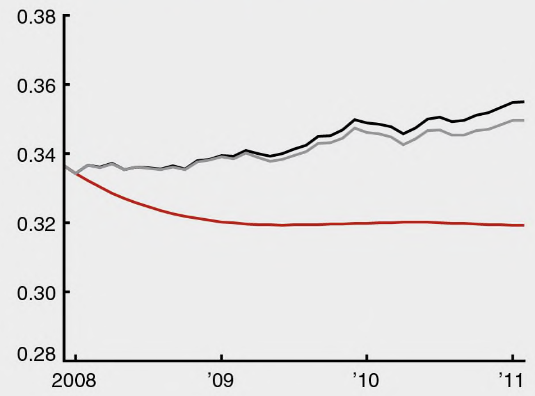
Effects on unemployment (3 states, transitionary dynamics)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 32

Effects on nonparticipation (3 states, transitionary dynamics)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

ψ_t can be interpreted as the search intensity of nonparticipant workers.

The transition rate from unemployment into employment λ_t^{UE} is given by:

$$42)\ \lambda_t^{UE} = \left(\frac{M_t}{U_t}\right)\left(\frac{U_t}{U_t+\psi_t N_t}\right) = \frac{A_t(U_t+\psi_t N_t)^{\alpha} V_t^{1-\alpha}}{U_t+\psi_t N_t},$$

since a fraction $\left(\frac{U_t}{U_t+\psi_t N_t}\right)$ of the total matches M_t is formed with unemployed workers. Similarly, the transition rate from nonparticipation into employment λ_t^{NE} is given by:

$$43)\ \lambda_t^{NE} = \left(\frac{M_t}{N_t}\right)\left(\frac{\psi_t N_t}{U_t+\psi_t N_t}\right) = \psi_t \frac{A_t(U_t+\psi_t N_t)^{\alpha} V_t^{1-\alpha}}{U_t+\psi_t N_t},$$

since a fraction $\left(\frac{\psi_t N_t}{U_t+\psi_t N_t}\right)$ of the total matches M_t is formed with nonparticipant workers.

The transition rate from employment to unemployment λ_t^{EU}, the transition rate from employment to nonparticipation λ_t^{EN}, the transition rate from unemployment to nonparticipation λ_t^{UN}, and the transition rate from nonparticipation to unemployment λ_t^{NU} are assumed to be exogenous to the model.

The evolution of workers across labor market states is then given by the following equations:

$$44)\ E_{t+1} = E_t + M_t - (\lambda_t^{EU} + \lambda_t^{EN})E_t,$$

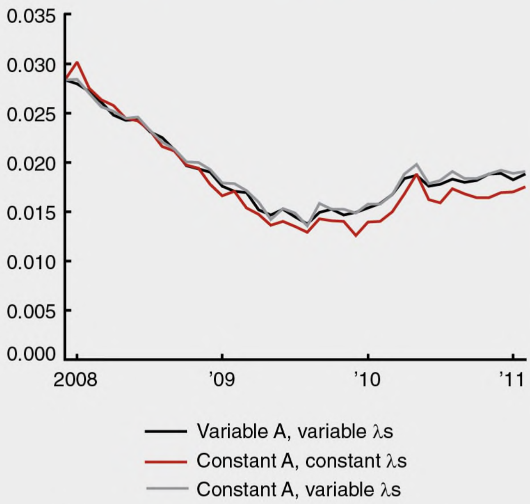
$$45)\ U_{t+1} = U_t + \lambda_t^{EU}E_t + \lambda_t^{NU}N_t - (\lambda_t^{UE} + \lambda_t^{UN})U_t,$$

$$46)\ N_{t+1} = N_t + \lambda_t^{EN}E_t + \lambda_t^{UN}U_t - (\lambda_t^{NE} + \lambda_t^{NU})N_t.$$

Equation 44 states that next-period employment is equal to current employment, plus all new matches, minus total separations (either to unemployment or nonparticipation). Equation 45 states that next-period unemployment is equal to current unemployment, plus all transitions into unemployment (either from employment or nonparticipation), minus all transitions out of unemployment (either to employment or nonparticipation). Equation 46 states that next-period nonparticipation is equal to current nonparticipation, plus all transitions into nonparticipation (either from employment or unemployment), minus all transitions out of nonparticipation (either to employment or unemployment).

FIGURE 33

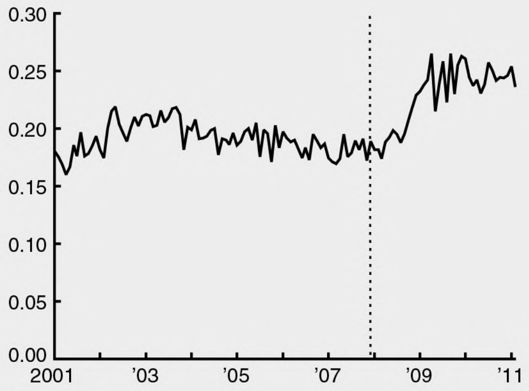
Effects on vacancies (3 states, transitionary dynamics)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 34

Search intensity of nonparticipants



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

The free-entry condition is given by

$$47)\ k = \left(\frac{M_t}{V_t}\right) J_t,$$

since from the point of view of a firm, the probability of filling a vacancy is now equal to M_t / V_t.

Observe that using equations 41, 42, and 43, we can rewrite equations 44–47 as follows:

$$48)\ E_{t+1} = E_t + A_t\left(U_t + \psi_t N_t\right)^{\alpha} V_t^{1-\alpha} - \left(\lambda_t^{EU} + \lambda_t^{EN}\right) E_t,$$

$$49)\ U_{t+1} = U_t + \lambda_t^{EU} E_t + \lambda_t^{NU} N_t - \left(\frac{A_t\left(U_t + \psi_t N_t\right)^{\alpha} V_t^{1-\alpha}}{U_t + \psi_t N_t} + \lambda_t^{UN}\right) U_t,$$

$$50)\ N_{t+1} = N_t + \lambda_t^{EN} E_t + \lambda_t^{UN} U_t - \psi_t\left(\frac{A_t\left(U_t + \psi_t N_t\right)^{\alpha} V_t^{1-\alpha}}{U_t + \psi_t N_t} + \lambda_t^{NU}\right) N_t,$$

$$51)\ V_t = \frac{J_t}{k} A_t\left(U_t + \psi_t N_t\right)^{\alpha} V_t^{1-\alpha}.$$

Given the total number of workers unemployed at date zero U_0 and the total number of workers that are nonparticipants at date zero N_0, the model generates an endogenous path for $\{M_t, V_t, U_{t+1}, N_{t+1}\}_{t=0}^{\infty}$.

Results

From equations 42 and 43, we have that the search intensity of nonparticipants can be measured as

$$52)\ \psi_t = \frac{\lambda_t^{NE}}{\lambda_t^{UE}}.$$

Figure 34 shows that the fraction of nonparticipants that search has increased quite substantially since the start of the latest recession.

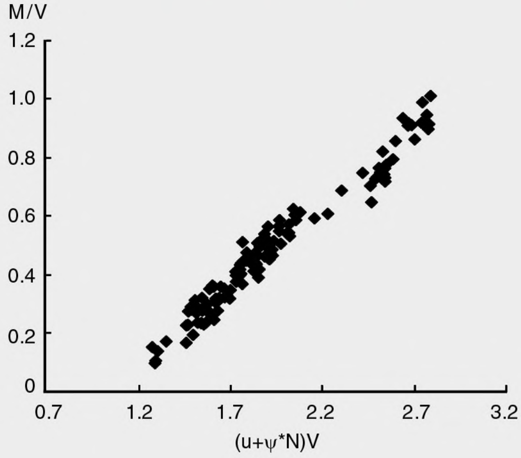
From equation 41, we have that

$$53)\ \ln\left(\frac{M_t}{V_t}\right) = \ln\left(A_t\right) + \alpha \ln\left(\frac{U_t + \psi_t N_t}{V_t}\right).$$

In what follows, I identify M_t with the total number of workers that transition from unemployment into employment between months t and $t + 1$, plus the total number of workers that transition from nonparticipation into employment between those same months,

FIGURE 35

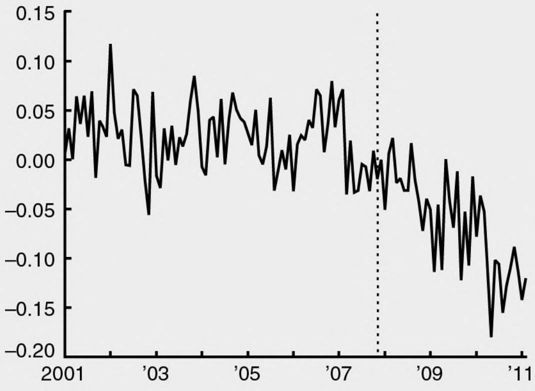
Matching function



Note: M/V is total matches per vacancy; (u+ψ*N)/V is total searchers per vacancy.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 36

Matching efficiency (transitionary dynamics, N in matching function)



Note: The log of A_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

as reported by the CPS. Figure 35 plots $\ln\left(\frac{M_t}{V_t}\right)$ against $\ln\left(\frac{U_t+\psi_t N_t}{V_t}\right)$ for the whole sample period. We see a strong linear relation, suggesting that equation 53 provides a good description of the data with a relatively constant A_t. Fitting equation 53 using OLS between 2001:1 and 2007:12 gives an estimate of $\alpha = 0.62$.

Given this estimated value of α, the path for the matching efficiency parameter A_t implied by equation 53 is reported in figure 36. We see that the path is not very different from figure 29 (p. 163). In turn, the value of a job, which is measured as

$$54)\ J_t = k\frac{V_t}{M_t},$$

is reported in figure 37. We also see that its path is not very different from figure 30 (p. 163).

In order to decompose the effects of the different variables on labor market dynamics, I find the E_t^*, U_t^*, N_t^*, and V_t^* that satisfy equations 48–51 under the assumption that $A_t = A_{2007:12}$, $\psi_t = \psi_{2007:12}$, $\lambda_t^{UN} = \lambda_{2007:12}^{UN}$, $\lambda_t^{EU} = \lambda_{2007:12}^{EU}$, $\lambda_t^{EN} = \lambda_{2007:12}^{EN}$, and $\lambda_t^{NU} = \lambda_{2007:12}^{NU}$, for every month t. That is, E_t^*, U_t^*, N_t^*, and V_t^* describe the employment, unemployment, nonparticipation, and vacancies levels that would have been obtained if J_t had changed the way it did but all other variables had remained constant at their December 2007 levels.

Also, I compute the E_t^{**}, U_t^{**}, N_t^{**}, and V_t^{**} that satisfy equations 48–51 under the assumption that $A_t = A_{2007:12}$ and $\psi_t = \psi_{2007:12}$, for every month t. That is, E_t^{**}, U_t^{**}, N_t^{**}, and V_t^{**} describe the employment, unemployment, nonparticipation, and vacancies levels that would have been obtained if J_t and all λ values had changed the way they did but A_t and ψ_t had remained constant at their December 2007 levels.

Similarly, I compute the E_t^{***}, U_t^{***}, N_t^{***}, and V_t^{***} that satisfy equations 48–51 under the assumption that $\psi_t = \psi_{2007:12}$, for every month t. That is, E_t^{***}, U_t^{***}, N_t^{***}, and V_t^{***} describe the employment, unemployment, nonparticipation, and vacancies levels that would have been obtained if A_t, J_t, and all λ values had changed the way they did but ψ_t had remained constant at its December 2007 level.

Figure 38 reports the paths for $U_t^* / (E_t^* + U_t^*)$ ("constant A, λ, ψ"), $U_t^{**} / (E_t^{**} + U_t^{**})$ ("constant A, ψ"), $U_t^{***} / (E_t^{***} + U_t^{***})$ ("constant ψ"), and $U_t / (E_t + U_t)$ ("everything variable"). From the $U_t^* / (E_t^* + U_t^*)$ path, we see that changes in the value of a job J_t played a very minor role in unemployment dynamics: From an initial unemployment rate of 4.9 percent, changes in J_t are only able to generate a peak unemployment rate of 6.4 percent. From comparing the path for $U_t^{**} / (E_t^{**} + U_t^{**})$ with the path for $U_t^* / (E_t^* + U_t^*)$, we see that changes in the transition rates λ played a crucial role in generating the large and persistent

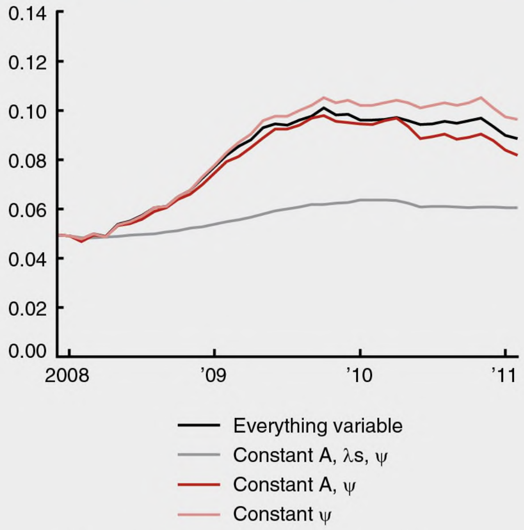
FIGURE 37

Value of a job (transitionary dynamics, N in matching function)

Note: The log of J_t is normalized to zero at the start of the past recession.
Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 38

Effects on unemployment (transitionary dynamics, N in matching function)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

increase in the unemployment rate. In fact, the path for $U_t^{**}/(E_t^{**}+U_t^{**})$ is very similar to $U_t/(E_t+U_t)$. Comparing $U_t^{***}/(E_t^{***}+U_t^{***})$ with $U_t^{**}/(E_t^{**}+U_t^{**})$, we see that the large drop in matching efficiency shown in figure 36 had a nontrivial role in the increase in the unemployment rate: The largest difference between $U_t^{***}/(E_t^{***}+U_t^{***})$ and $U_t^{**}/(E_t^{**}+U_t^{**})$ is 1.5 percent. Comparing $U_t^{***}/(E_t^{***}+U_t^{***})$ with $U_t/(E_t+U_t)$, we see that the increase in the search intensity of nonparticipants shown in figure 34 roughly offsets the effects of the fall in matching efficiency.

Figure 39 reports the paths for N_t^* ("constant A, λ, ψ"), N_t^{**} ("constant A, ψ"), N_t^{***} ("constant ψ") and N_t ("everything variable"). We see that the value of a job had a significant effect on nonparticipation: The gray line first declines and then increases quite rapidly. The changes in transition rates first increased nonparticipation but then lowered it: The red line is initially above the gray line, but it crosses it in mid-2009. The drop in matching efficiency of figure 36 has the effect of increasing nonparticipation: The light red line is significantly higher than the red line. However, the large increase in the search intensity of nonparticipants shown in figure 34 had a large effect on reducing nonparticipation: The black line is much lower than the light red line.

Lastly, figure 40 shows that the only important determinant of vacancy dynamics was the value of a job J_t (all other lines are quite similar to the gray).

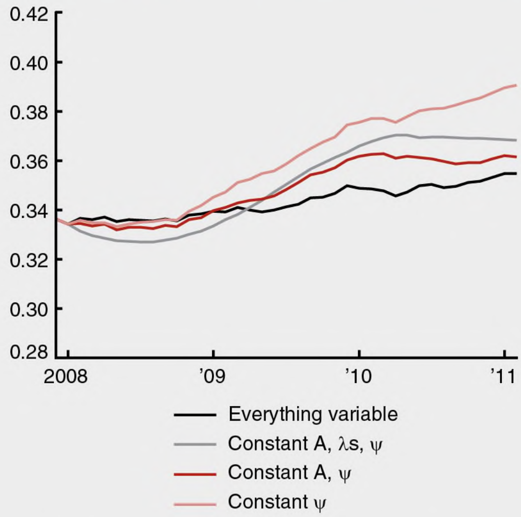
Conclusion

This article has explored different approaches to measuring matching efficiency and assessing its implications for labor market dynamics since the start of the past recession. In particular, I evaluated the importance of allowing for a third labor market state, allowing for variable transition rates, considering explicit transitionary dynamics, and allowing nonparticipants to enter the matching function. I find that the results are quite sensitive to the different specifications.

In the preferred scenario (that is, the case with three labor market states, variable transition rates, nonparticipants entering the matching function, and explicit transitionary dynamics), I obtained the following findings. First, the matching efficiency parameter is quite volatile throughout the sample period. Second, the matching efficiency has been drifting down since the start of the recession. Third, the value of filled jobs plummeted between 2007:12 and 2009:6 but has recovered quite significantly since then. Fourth, conditional on the observed paths for the value of a job and all transition rates, the drop in matching efficiency since the start of the recession has had only a moderate im-

FIGURE 39

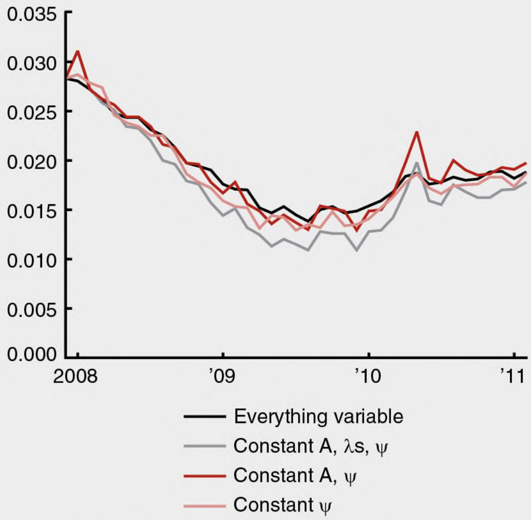
Effects on nonparticipation (transitionary dynamics, N in matching function)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

FIGURE 40

Effects on vacancies (transitionary dynamics, N in matching function)



Sources: Author's calculations based on data from the Bureau of Labor Statistics' *Current Population Survey* and *Job Openings and Labor Turnover Survey*.

pact on the unemployment rate. Fifth, the large effects on unemployment rate dynamics arise from changes in the transition rates. Sixth, the matching efficiency, the value of a job, the transition rates, and the search intensity of nonparticipants all have significant effects on the dynamics of nonparticipation.

The analysis performed in this article decomposed the observed growth in unemployment and nonparticipation into contributions from changes in matching efficiency, the value of a job, the search intensity of nonparticipants, and the transition rates across different labor market states. This decomposition was done very much in the spirit of standard growth accounting exercises, in which GDP growth is decomposed into growth contributions from total factor productivity, capital, and labor. Interpreted as a growth accounting exercise, the results in this paper should be considered as extremely informative. However, care should be exercised in providing a counterfactual interpretation to the results. The reason is that if the matching efficiency had stayed constant at its 2007:12 level (instead of dropping as it actually did), this would have affected the value of a job and the transition rates across labor market states, but these secondary effects have not been considered in the analysis. That is, the contributions to labor market dynamics of the value of a job, the productivity of the matching function, the different transition rates, and the search intensity of nonparticipants have been calculated as not affecting the other variables.[10] To evaluate counterfactuals such as this, the explicit economic decisions and wage determination process that this paper has abstracted from would have to be incorporated and the equilibrium of such full-blown models would have to be analyzed. Of course, the counterfactual results would depend on how those modeling choices are made.

Another caveat to the analysis is that it has not incorporated job-to-job transitions. Since the workers making these transitions are competing for the same pool of vacancies as unemployed and nonparticipant workers, their behavior may affect the measurement of matching efficiency.

NOTES

[1]The negative relation between unemployment and vacancies is not exclusive to the U.S.: It is present in a number of countries. See Petrongolo and Pissarides (2001) for an empirical survey of the Beveridge curve and the matching function.

[2]Strictly speaking, what these papers implicitly assume is that the transitionary dynamics of the model economy are extremely fast. Under this assumption, they use the steady state of the model to analyze data, even when the matching function and transition rates change over short periods.

[3]The preferred specification is the one with the least restrictive and/or most appealing assumptions.

[4]We could choose alternative approaches. For instance, we could take the average of the ratios in equation 11 for every pair of months (i, j). The problem with this approach is that measurement errors would be severely amplified for pairs of months with similar unemployment/vacancy ratios. Another approach would be to take the average of the ratios in equation 11 for pairs of months with sufficiently large differences in unemployment/vacancy ratios. This approach would lead to a similar estimate for α as the approach chosen here.

[5]For U_t, I use the unemployment rate at month t from the Bureau of Labor Statistics' (BLS) *Current Population Survey* (CPS). For V_t, I use the average vacancies reported by the BLS's *Job Openings and Labor Turnover Survey* (JOLTS) in months t and $t-1$, divided by the size of the labor force reported by the CPS in month t. I average the vacancies reported by JOLTS because they correspond to the number of vacancies at the end of the month, while the CPS data roughly correspond to observations in the middle of the month.

[6]This is a large number. However, J_t must be interpreted as the value of creating an additional job and not as the average value of all existing jobs. Obtaining such a large drop in the value of a marginal job should not be surprising, given the severity of the recession experienced by the U.S. at the time.

[7]See the conclusions for a further discussion of how to interpret the results of this article.

[8]See note 2 for a discussion of this estimation approach.

[9]These results should not be taken seriously. In what follows, I show that the large drop in matching efficiency obtained in figure 16 is an artifact of the constant transition rates.

[10]Observe that similar caution must be used in providing a counterfactual interpretation to standard GDP growth accounting exercises. If total factor productivity had stayed constant during the last 50 years (instead of growing at its actual rate), this would have affected capital accumulation and labor supply. However, these secondary effects are not taken into account when reporting the contribution to growth of total factor productivity.

REFERENCES

Barlevy, Gadi, 2011, "Evaluating the role of labor market mismatch in rising unemployment," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 35, Third Quarter, pp. 82–96.

Barnichon, Regis, and Andrew Figura, 2010, "What drives movements in the unemployment rate? A decomposition of the Beveridge curve," Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System, working paper, No. 2010-48.

Kocherlakota, Narayana, 2010, "Back inside the FOMC," speech in Missoula, MT, September 8.

Mortensen, Dale T., and Christopher A. Pissarides, 1994, "Job creation and job destruction in the theory of unemployment," *Review of Economic Studies*, Vol. 61, No. 3, July, pp. 397–415.

Petrongolo, Barbara, and Christopher A. Pissarides, 2001, "Looking into the black box: A survey of the matching function," *Journal of Economic Literature*, Vol. 39, No. 2, June, pp. 390–431.

Shimer, Robert, 2005, "The cyclicality of hires, separations, and job-to-job transitions," *Review*, Federal Reserve Bank of St. Louis, Vol. 87, No. 4, July/August, pp. 493–507.