Economic Quarterly—Volume 99, Number 1—First Quarter 2013—Pages 1–23

Why Labor Force Participation (Usually) Increases when Unemployment Declines

Andreas Hornstein

During the Great Recession, the unemployment rate increased rapidly within two years from about 4 percent in 2007 to about 10 percent in 2009. Yet over the ensuing recovery, the unemployment rate has declined only gradually and, more than four years after the end of the recession, it now stands at about 7 percent. At the same time, the labor force participation rate has declined steadily over this time period and now stands at about 63 percent, a level comparable to the early 1980s. Many observers view the decline in the labor force participation rate as an indication that further declines in the unemployment rate will come only slowly. The expectation is that if the labor market improves, many participants who have left the labor market will return and contribute to the pool of unemployed, and many unemployed participants will no longer exit the labor force but continue to search for work.[1]

Past business cycles have indeed been characterized by a negative correlation between the unemployment rate and the labor force participation (LFP) rate, that is, as the unemployment rate declines, the LFP rate increases. In this article we use observations on gross flows

This is a revised version of an article previously titled "The Cyclicality of the Labor Force Participation Rate." I would like to thank Marianna Kudlyak, John Muth, Felipe Schwartzman, and Alex Wolman for helpful comments. Any opinions expressed are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: andreas.hornstein@rich.frb.org.

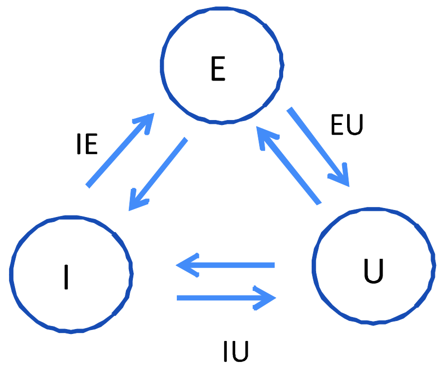
[1] For example, see Daly et al. (2012), Hatzius (2012), Davidson (2013), or Tankersley (2013).

between labor market states to provide a more detailed analysis of why the unemployment rate and the LFP rate are negatively correlated over the business cycle. For our analysis, the total potential workforce is decomposed into three groups: the employed (E), the unemployed (U), and the out-of-the-labor-force group, or inactive (I) for short. The LFP rate is the share of employed and unemployed in the potential workforce, and the unemployment rate is the share of the unemployed in the labor force. We think of labor market participants as transitioning between these three states. Figure 1 provides a stylized representation of these transitions. The arrows connecting the circles represent the gross flows between the three labor market states. For our analysis we look at a gross flow as the product of two terms: the total number of participants that could potentially make a transition and the rate at which the participants make the transition. For example, the total number of unemployed who become employed is the product of the number of unemployed and the probability at which an unemployed worker will become employed. The transition probabilities reflect the opportunities faced and choices made by labor market participants. For example, the probability of an unemployed worker becoming employed depends, among other things, on the number of available jobs (vacancies) and the search effort while unemployed. Given the size of the potential workforce, the transition rates between labor market states determine the LFP rate and the unemployment rate.

We have marked three groups among the transitions in Figure 1: EU, IU, and IE. The first group involves transitions within the labor force, between employment and unemployment, and these transitions have been the focus of much recent research on the determination of the unemployment rate.[2] The working assumption of this research has been that, for an analysis of the unemployment rate, a fixed LFP rate is a reasonable first approximation. The second and third group involve transitions between the labor force and out-of-the-labor-force, that is, they potentially generate changes of the LFP rate. The second group, which involves transitions between inactivity and unemployment, is at the heart of the above mentioned concern that further reductions in the unemployment rate will come only slowly. This concern is based on the assumption that, as the labor market improves, unemployed workers become less likely to exit the labor force and inactive workers become more likely to join the labor force as unemployed; we call this the IU hypothesis.

[2] For example, see Shimer (2012) and other research mentioned below.

Figure 1 Labor Market State Transitions



In this article we argue that observations on transition probabilities obtained from gross flow data are inconsistent with the IU hypothesis. In fact, the opposite is true: As the labor market improves, unemployed workers become more likely to exit the labor force and inactive workers become less likely to join the labor force as unemployed. This pattern for IU transitions would result in a positive correlation between the unemployment rate and the LFP rate. The observed negative correlation between unemployment and LFP must then result from patterns in the EU and IE group transition rates. We calculate the contributions of cylical variations in the transition rates for the three groups—IU, IE, and EU—and indeed find that the variations in the IE and EU group transition rates generate a negative co-movement of the unemployment and LFP rates that dominates the positive co-movement generated by the IU group transition rates. This suggests that an increasing LFP rate is more the by-product of an improving labor market rather than a brake on the declining unemployment rate.

This article is based on a line of research that accounts for changes in labor market ratios through changes in the rates at which labor market participants transition between labor market states. Early work in this literature mostly ignored variations in the LFP rate and focused on variations in transition rates between the two labor market states—employment and unemployment—for example, Elsby, Michaels, and

Solon (2009), Fujita and Ramey (2009), and Shimer (2012). This work finds that variations in unemployment exit rates contribute relatively more to unemployment rate volatility than do variations in employment exit rates. Recently, a similar approach has been applied to a more general accounting framework that adds a third labor market state, out-of-the-labor-force, and allows for variations in the LFP rate, for example, Barnichon and Figura (2010) and Elsby, Hobijn, and Şahin (2013).[3] Our work is closest to Elsby, Hobijn, and Şahin (2013), but their main focus is on accounting for the relative contributions of transition rate volatility to unemployment rate volatility.[4] Nevertheless, they also point out that the cyclical behavior of measured transition rates between unemployment and inactivity is at odds with common preconceptions about that behavior, and they also note that the observed cyclical behavior of these transition rates would induce a positive correlation between the unemployment rate and the LFP rate.

The article is organized as follows. Section 1 documents the negative correlation between the detrended unemployment rate and LFP rate for the total working age population, and men and women separately. Section 2 documents the co-movements between the unemployment rate and transition probabilities between labor market states. Section 3 demonstrates how variations in transition rates contribute to the co-movement of the unemployment rate and the LFP rate. In conclusion, Section 4 speculates on the implications of the recent "unusual" co-movement of unemployment and LFP in the recovery since 2010.

1. UNEMPLOYMENT AND LFP

The U.S. Bureau of Labor Statistics (BLS) publishes monthly data on the labor market status of U.S. households that are based on the Current Population Survey (CPS). The CPS surveys about 60,000 households every month with about 110,000 household members, a representative sample of the U.S. working age population. Household respondents are asked if the household members are employed, and if

[3] Shimer (2012) also develops tools for the analysis of a multi-state labor market model and studies the role of variations in the LFP rate, but the focus of the article is on the two-state model of the labor market.

[4] An important part of Elsby, Hobijn, and Şahin (2013) is their analysis of a measurement issue for gross flows. Since gross flows are derived from survey samples, it is always possible that survey respondents are misclassified with respect to their labor market state. Past research has demonstrated that misclassification is a significant issue. Elsby, Hobijn, and Şahin (2013) argue that allowing for the possibility of misclassification does not substantially affect the conclusions drawn from measured gross flows for the issue studied in this article.

they are not employed, whether they want to work and are actively looking for work. The latter are considered to be unemployed, and employed and unemployed household members constitute the labor force. Household members that are not employed and that are not actively looking for work are considered to be not part of the labor force, or inactive for short. The unemployment rate is the share of unemployed workers in the labor force, and the LFP rate is the share of the labor force in the working age population.[5]

The unemployment rate tends to be more volatile than the LFP rate in the short run, but changes in the LFP rate tend to be more persistent over the long run. Figure 2, panels A and B, display quarterly averages of monthly unemployment and LFP rates for the period from 1948 to 2012. The unemployment rate increases sharply in a recession, and then declines gradually during the recovery. Shaded areas in Figure 2 indicate periods when the unemployment rate is increasing, and these periods match periods of National Bureau of Economic Research (NBER) recessions quite well.[6] Even though the average unemployment rate appears to be somewhat higher than usual in the 1970s, considering the magnitude of short-run fluctuations in the unemployment rate, the average unemployment rate does not change much over subsamples of the period. The 2007–09 Great Recession stands apart by the magnitude of the increase of the unemployment rate and the rather slow decline of the unemployment rate from its peak.

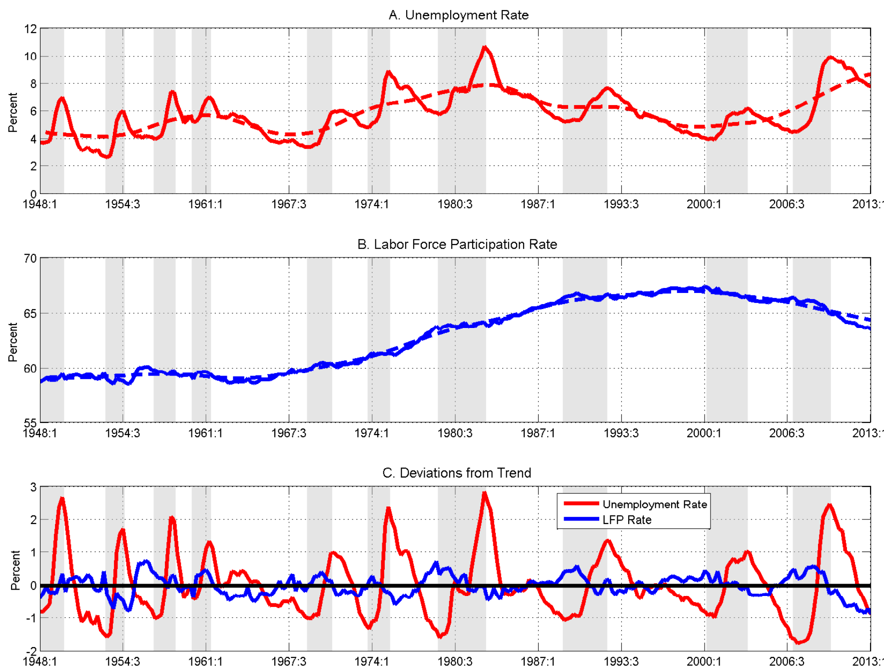
The LFP rate does not display much short-run volatility, rather it is dominated by long-run demographic trends. Starting in the mid-1960s, the LFP rate increased gradually from values slightly below 60 percent to reach a peak of 67 percent in 2000. This slow but persistent increase of the LFP rate can be accounted for by the increasing LFP rate of women and early on by the baby boomer generation entering the labor force. Since 2000, the LFP rate has declined, first gradually, then at an accelerated rate since the Great Recession and is now at about 63 percent. The gradual decline in the LFP rate can be attributed to the aging of the baby boomer generation and declining LFP rates for women and the young (less than 25 years of age).[7] In general, there is not much short-run volatility in the LFP rate, the recent accelerated

[5] Households are asked about other features of their labor market status, but the questions about employment and active search for work when not employed are the main questions of interest for determining the unemployment rate and the LFP rate. For a detailed description of the survey and the methods used, see Bureau of Labor Statistics (2012).

[6] The business cycle dates provided by the NBER are a widely accepted measure of the peaks and troughs of U.S. economic activity.

[7] For example, see Aaronson et al. (2006).

Figure 2 Unemployment and Labor Force Participation, 1948–2013



Notes: The unemployment and LFP rates displayed in panels A and B are quarterly averages of monthly values. Shaded (white) areas are periods when the unemployment rate is increasing (declining). The dashed lines are the trend calculated using a Baxter and King (1999) bandpass filter series with periodicity more than 12 years for the trend. Panel C displays the difference between actual and trend values of the unemployment rate and the LFP rate.

decline following the Great Recession being the exception. This accelerated decline in the LFP rate after the Great Recession shows up in the declining LFP rates of mature workers between 25 and 55 years of age, especially men, and also in declining participation rates of the young.

Table 1 Cyclicality of Unemployment and Labor Force Participation

Sample	σ_u	σ_l	Corr($u(t)$,$l(t+s)$) for s=								
			−4	−3	−2	−1	0	1	2	3	4
					Total						
1952:Q1–2007:Q4	0.89	0.29	−0.09	−0.20	−0.30	−0.38	−0.45	−0.52	−0.55	−0.54	−0.48
1952:Q1–1991:Q4	0.93	0.31	−0.09	−0.19	−0.29	−0.37	−0.43	−0.49	−0.53	−0.51	−0.44
1992:Q1–2007:Q4	0.79	0.21	−0.08	−0.21	−0.39	−0.55	−0.65	−0.71	−0.70	−0.69	−0.68
1992:Q1–2013:Q1	0.98	0.33	0.08	−0.07	−0.24	−0.41	−0.53	−0.63	−0.70	−0.75	−0.75
					Men						
1952:Q1–2007:Q4	1.01	0.28	−0.03	−0.18	−0.30	−0.39	−0.45	−0.52	−0.55	−0.55	−0.48
1952:Q1–1991:Q4	1.04	0.28	−0.09	−0.22	−0.34	−0.41	−0.46	−0.52	−0.55	−0.53	−0.44
1992:Q1–2007:Q4	0.92	0.27	0.14	−0.03	−0.27	−0.48	−0.61	−0.70	−0.74	−0.77	−0.77
1992:Q1–2013:Q1	1.19	0.41	0.07	−0.09	−0.27	−0.45	−0.57	−0.67	−0.73	−0.78	−0.78
					Women						
1952:Q1–2007:Q4	0.77	0.36	−0.16	−0.22	−0.28	−0.34	−0.37	−0.42	−0.45	−0.43	−0.37
1952:Q1–1991:Q4	0.81	0.40	−0.13	−0.20	−0.25	−0.32	−0.35	−0.41	−0.45	−0.43	−0.36
1992:Q1–2007:Q4	0.65	0.23	−0.26	−0.30	−0.38	−0.43	−0.43	−0.46	−0.38	−0.35	−0.31
1992:Q1–2013:Q1	0.77	0.32	0.07	−0.04	−0.17	−0.29	−0.39	−0.49	−0.58	−0.63	−0.64

Notes: Standard deviations and cross-correlations of detrended unemployment, u, and labor force participation rate, l, for total, men, and women. The trend for each variable is calculated as a Baxter and King (1999) bandpass filter with periodicity more than 12 years for monthly data, from January 1948 to March 2013. Unemployment and LFP rate are in percent, and detrended values are the difference between actual values and trend. Statistics are calculated for quarterly averages of monthly data for the indicated subsamples.

The average unemployment rate in the 1960s, when the LFP rate was low, does not appear to be much different from the average unemployment rate in the 1990s when the LFP rate was high. In other words, the unemployment rate and the LFP rate do not appear to be correlated over the long run. Over the short run, the unemployment rate and the LFP rate are, however, negatively correlated, that is, the LFP rate increases as the unemployment rate declines.

We define short-run movements of the unemployment rate and the LFP rate as deviations from trend, and we define the trend of a time series as a smooth line drawn through the actual time series. To be precise, we construct the trend using a bandpass filter that extracts movements with a periodicity of more than 12 years.[8] The dashed lines in Figure 2, panels A and B, display the trends for the unemployment rate and the LFP rate.[9] In panel C of Figure 2 we display the deviations from trend, that is, the difference between the actual and trend values, for the LFP rate and the unemployment rate. Clearly, deviations from trend are more volatile for the unemployment rate than for the LFP rate. Furthermore, the LFP rate tends to be above trend whenever the unemployment rate is below trend and vice versa. In Table 1 we display the standard deviations and cross-correlations between the detrended unemployment rate and the LFP rate for the total working age population, and for men and women separately.

The unemployment rate is about three times as volatile as the LFP rate, and the LFP rate increases as the unemployment rate declines, with the LFP rate lagging about half a year.[10] When we split the sample in the early 1990s, we can see that both the unemployment rate and the LFP rate are less volatile since the 1990s, but they remain negatively correlated.[11] Including the Great Recession and its

[8] We use the method of Baxter and King (1999) to construct the trend. This is just one of several alternative methods to calculate trends. The results do not differ much if instead we use a Hodrick and Prescott (1997) filter, or a random walk bandpass filter as described in Christiano and Fitzgerald (2003).

[9] At the beginning and end of the sample, our procedure delivers an ill-defined measure of the trend. Essentially, the trend of a series is a symmetric moving average of the series. Thus, at the beginning and end of the sample, we do not have enough data points to calculate the trend. For these truncated periods we simply choose to truncate the moving average filter and reweigh the available data points. This procedure is arbitrary, and it implies that current data points receive much more weight in determining the trend, which explains the high trend value for the unemployment rate in 2012. For the statistical analysis below we therefore discard some observations at the beginning and end of sample, and start the sample in 1952:Q1 and end the sample in 2007:Q4.

[10] We define the length of the lead/lag by the correlation that is largest in absolute value.

[11] This is consistent with the period being part of the "Great Moderation" in the United States, which indicates an economy-wide decline in volatility starting in the mid-1980s. We choose to split the sample in 1992 because in the next section we study how changes in labor market transition rates contribute to the co-movement of the

aftermath significantly increases the measured volatility of the unemployment rate and LFP rate, but, again, it does not much affect the measured negative correlation between the two variables.[12] Finally, the cyclical co-movement between unemployment and LFP is similar for men and women, but the unemployment rate is relatively more volatile for men, the LFP rate is relatively more volatile for women, and the LFP rate is lagging the unemployment rate more for men than for women.

We now study if this negative correlation between the unemployment rate and the LFP rate can be accounted for by inactive workers becoming more likely to enter the labor force and unemployed workers becoming less likely to exit the labor force.

2. TRANSITIONS BETWEEN LABOR MARKET STATES

The CPS household survey not only contains information on how many people are employed, unemployed, and inactive in any month, but it also contains information on how many people switch labor market states from one month to the next. We can use these gross flows between labor market states to calculate the probabilities that any one household member will, within a month, transition from one labor market state to a different state. This information can be used to see if, for example, variations in the transition rates between inactivity and unemployment are consistent with the usual interpretation of the negative co-movement of the unemployment rate and the LFP rate.

Households are surveyed repeatedly in the CPS. In particular, the survey consists of a rotation sample, that is, once a household enters the sample it is surveyed for four consecutive months, then it leaves the sample for eight months, after which it reenters the sample and is once more surveyed for four consecutive months. Thus, in any month, for three-fourths of the household members in the sample, we potentially have observations on their current labor market state and their state in the previous month. We can use this information to calculate the gross flows between labor market states from one month to the

unemployment rate and the LFP rate. We calculate transition rates from data on gross flows for the period after 1990, and again we discard some of the beginning and end of sample data on deviations from trend to minimize the problems arising from an ill-defined trend.

[12] Related to the discussion in footnote 9, we should note that if the unemployment rate continues to decline, then future measures of the trend unemployment rate that include these data points will indicate a lower trend unemployment rate than do our current measures. Thus, our current measure very likely understates the cyclical deviations from trend for the unemployment rate.

next. The measurement of gross flows suffers from two problems, missing data points and misclassified data points. We will use data series for gross flows that have been adjusted for missing data but not for misclassification.[13]

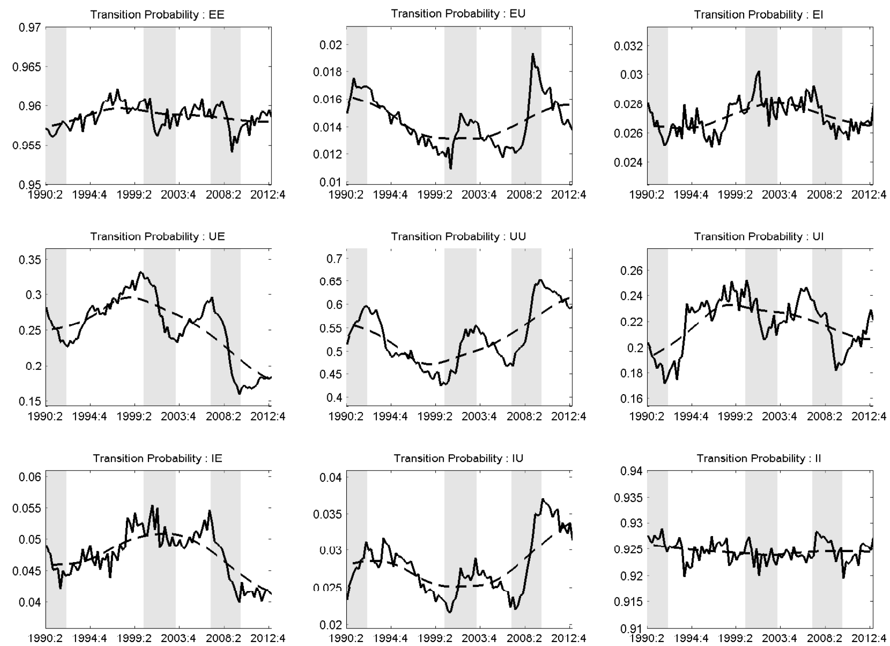
Data points are missing because the actual unit of observation in the CPS is not a particular household, but the household that is residing at a particular address. Thus, even for those addresses that have entered the sample in the previous month, we may not have observations on the previous month's labor market states for the members of the current resident household. This might happen for various reasons. The household could have a new member who did not live at the current address in the previous month, for example, a dependent returning to the family household after a longer absence. Alternatively, the household previously residing at the address moved away and a new household moved in. About 15 percent of the potential observations cannot be matched across months, and these observations are not missing at random (Abowd and Zellner 1985). One can use "margin adjustment" procedures to generate gross flow data consistent with unconditional marginal distributions, and these procedures take into account the possibility that observations are not missing at random. In the following, we use the BLS-provided margin adjusted research series on labor force status flows from the CPS.[14]

Gross flows from one labor market state to another can be interpreted as the product of two terms: the total number of participants in the initial state and the probability that any one of these participants makes the transition from the initial state to another state. For example, more people might make the transition from unemployment to inactivity because there are more unemployed people, or because each unemployed worker is more likely to make the transition. In Figure 3 we display the transition probabilities between employment (E), unemployment (U), and inactivity (I) that are implied by the observed gross flows between labor market states for the period from 1990 to 2012. A panel labeled AB denotes the probability that a participant who is in labor market state A will transition to state B within a month. For example, the center panel in the bottom row, labeled IU, denotes the probability that a participant who is inactive in the current month will

[13] The evidence for misclassification in the BLS, that is, that a participant is assigned the wrong labor market state in the survey, has been discussed for a long time, see, for example, Poterba and Summers (1986). There is currently no generally accepted procedure to adjust CPS data on labor market states for misclassification. Recently, Elsby, Hobijn, and Şahin (2013) and Feng and Hu (2013) have worked on possible corrections for misclassification.

[14] The research series is available at www.bls.gov/cps/cps_flows.htm. Frazis et al. (2005) describe the BLS procedure used to construct the series.

Figure 3 Transition Probabilities, 1990:Q2–2013:Q1



Notes: Panel AB denotes the probability of making the transition from labor market state A to labor market state B. The dashed lines are the trend calculated using a Baxter and King (1999) bandpass filter series with periodicity more than 12 years for the trend. The probabilities displayed are quarterly averages of monthly values. Shaded (white) areas are periods when the unemployment rate is increasing (declining).

be unemployed in the next month. Regions that are (not) shaded denote periods when the unemployment rate increases (declines). The trend for each transition probability is calculated using the same bandpass filter as in the previous section, and it is displayed as a dashed line in Figure 3. In Table 2, we display the average transition probabilities, the standard deviations of the detrended transition probabilities, and their cross-correlations with the detrended unemployment rate for the total working age population, and for men and women separately.

An increase in the unemployment rate is associated with more churning in the labor market: Employed workers are more likely to

Table 2 Cyclicality of Transition Probabilities

	$\bar{p}_{ij}$	σ_{ij}	Corr($u(t), p_{ij}(t+s)$) for s=								
			−4	−3	−2	−1	0	1	2	3	4
Total, $\bar{u} = 5.3$, $\sigma_u = 0.76$											
EU	1.4	0.10	0.70	0.83	0.88	0.88	0.85	0.72	0.62	0.51	0.42
UE	27.5	2.35	−0.48	−0.64	−0.78	−0.89	−0.95	−0.94	−0.88	−0.78	−0.65
IU	2.6	0.21	0.36	0.49	0.61	0.71	0.79	0.78	0.77	0.75	0.70
UI	22.4	1.39	−0.59	−0.68	−0.75	−0.79	−0.77	−0.68	−0.55	−0.36	−0.16
IE	4.9	0.21	−0.24	−0.35	−0.50	−0.57	−0.65	−0.66	−0.60	−0.55	−0.45
EI	2.7	0.09	−0.02	−0.02	−0.10	−0.24	−0.32	−0.45	−0.48	−0.45	−0.36
Men, $\bar{u} = 5.4$, $\sigma_u = 0.88$											
EU	1.5	0.13	0.73	0.85	0.89	0.90	0.86	0.73	0.63	0.53	0.43
UE	29.0	2.54	−0.46	−0.62	−0.76	−0.86	−0.92	−0.91	−0.85	−0.77	−0.65
IU	3.2	0.30	0.47	0.56	0.66	0.76	0.84	0.79	0.76	0.72	0.68
UI	18.9	1.47	−0.54	−0.62	−0.70	−0.77	−0.77	−0.71	−0.59	−0.41	−0.17
IE	5.7	0.27	−0.20	−0.33	−0.45	−0.53	−0.58	−0.62	−0.58	−0.50	−0.43
EI	2.2	0.07	−0.03	0.08	0.09	0.03	−0.00	−0.16	−0.19	−0.23	−0.20
Women, $\bar{u} = 5.3, \sigma_u = 0.63$											
EU	1.2	0.07	0.39	0.57	0.67	0.68	0.70	0.57	0.48	0.40	0.34
UE	25.8	2.31	−0.50	−0.62	−0.77	−0.86	−0.91	−0.90	−0.84	−0.73	−0.59
IU	2.3	0.18	0.21	0.35	0.48	0.60	0.71	0.68	0.69	0.68	0.61
UI	26.7	1.30	−0.54	−0.62	−0.68	−0.68	−0.66	−0.53	−0.40	−0.22	−0.08
IE	4.5	0.21	−0.21	−0.32	−0.46	−0.48	−0.61	−0.60	−0.53	−0.51	−0.39
EI	3.4	0.14	−0.03	−0.08	−0.18	−0.34	−0.43	−0.53	−0.54	−0.47	−0.36

Notes: The first column lists the sample average for transition probabilities from labor market state i to j, p_{ij}, with labor market states being employed (E), unemployed (U), and out-of-the-labor-force/inactive (I). The second column lists standard deviations of detrended transition probabilities, and the remaining columns list cross-correlations of detrended transition probabilities with the detrended unemployment rate. The trend for each variable is calculated as a Baxter and King (1999) bandpass filter with periodicity of more than 12 years for monthly data, from January 1990 to March 2013. Transition probabilities and the unemployment rate are in percent, and detrended values are the difference between actual and trend values. Statistics are calculated for quarterly averages of monthly data for the sample 1992:Q1 to 2007:Q4.

lose their jobs, and unemployed workers are less likely to return to work, with job loss (finding) rates slightly leading (lagging) the unemployment rate; see the panels labeled EU and UE in Figure 3 and the corresponding correlations in Table 2.[15] Considering the magnitude and volatility of the job finding rate for unemployed workers, the transition rate UE, it is apparent that variations in this rate are a

[15] In fact, when unemployment is high, gross flows between unemployment and employment are both high. Despite the lower probability of the unemployed finding employment, gross flows from unemployment to employment are high because there are more unemployed.

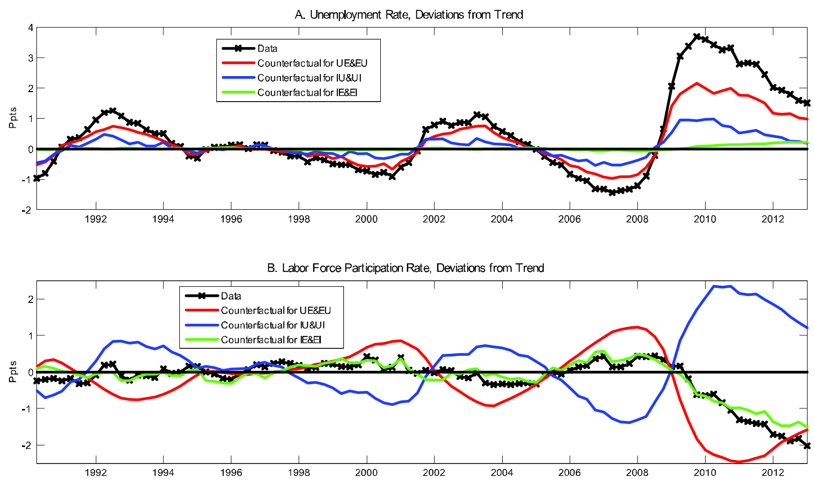
major source of unemployment volatility. Looking at panels IU and UI, we can see that as the unemployment rate declines, it becomes more likely that an unemployed worker exits the labor force and less likely that an inactive worker joins the labor force as unemployed. This pattern is confirmed by the cross-correlations for the detrended rates in Table 2. Thus, the cyclical pattern of the transition rates between inactivity and unemployment is exactly the opposite of what the IU hypothesis proposes as an explanation of the negative correlation between the LFP rate and the unemployment rate. However, the transition probabilities between inactivity and employment do have a cyclical pattern that supports a negative co-movement between the unemployment rate and the LFP rate. As the unemployment rate increases it becomes less likely that people make the transition from inactivity to employment. It also becomes less likely that employed workers leave the labor force, but this probability is always quite low and it is not very volatile over the cycle. The cyclical properties of the transition probabilities for all three groups, EU, IU, and IE, are roughly the same for men and women. The only exception is that transition probabilities for women tend to be somewhat less volatile overall, and that men's transition probabilities from employment to inactivity appear to be acyclical.

So far we have shown that the direct evidence on labor market transitions does not support the IU hypothesis of why the LFP rate increases as the unemployment rate declines. In particular, as the labor market improves and the unemployment rate declines, participants become less likely to make the transition from inactivity to unemployment and they become more likely to make the transition from unemployment to inactivity. So what accounts for the negative correlation of unemployment and the LFP rate?

3. SOURCES OF CO-MOVEMENT

Recent research on labor markets using the stock-flow approach points to the importance of variations in the job finding rate and job loss rate for the determination of the unemployment rate. We now argue that variations in the job finding and job loss rates are also important for the cyclical co-movement between the unemployment and LFP rates. As a first step, note that the exit rate from the labor force is an order of magnitude smaller for employed workers than it is for unemployed workers (see Table 2). This means that as the unemployment rate declines, the average exit rate from the labor force declines, and the LFP rate increases. Furthermore, as we have just seen, when the unemployment rate declines, more people join the labor force without

Figure 4 Counterfactuals for Unemployment Rate and LFP Rate



an intervening unemployment spell. This suggests that cyclical movements of the transition rates in the UE and IE group account for the negative co-movement of unemployment and LFP over the business cycle. We now formalize this argument by constructing counterfactuals for the unemployment rate and the LFP rate.

Consider the trend paths for the transition probabilities that we have calculated for Figure 3 and Table 2. We can interpret the deviations of the unemployment rate and the LFP rate from their respective trends as arising from deviations of the transition probabilities from their respective trends. In the Appendix, we describe a procedure that allows us to decompose the cyclical movements of the unemployment and LFP rates into parts that originate from the cyclical movements of the various transition probabilities.[16] In Figure 4, we graph the contributions to trend deviations of the unemployment rate and LFP rate (black lines) coming from variations in the transition probabilities between (1) employment and unemployment (red lines), (2) inactivity and unemployment (blue lines), and (3) inactivity and employment (green

[16] The procedure used to derive the contributions coming from variations in month-to-month transition probabilities is actually based on a model that allows for continuous transitions between labor market states in between the monthly survey dates.

Table 3 Cross-Correlations between Unemployment Rate and LFP Rate for Counterfactuals, Deviations from Trend, 1992:Q1–2007:Q4

	Corr(u(t), l(t+s)) for s=								
	−4	−3	−2	−1	0	1	2	3	4
UE and EU	−0.20	−0.40	−0.58	−0.74	−0.87	−0.95	−0.99	−0.97	−0.91
IU and UI	0.15	0.31	0.48	0.64	0.82	0.89	0.92	0.90	0.84
UE, EU, UI, and IU	0.41	0.37	0.32	0.24	0.23	0.13	0.04	−0.02	−0.07
IE and EI	−0.33	−0.50	−0.66	−0.86	−0.99	−0.83	−0.65	−0.55	−0.43
Actual	−0.10	−0.22	−0.40	−0.55	−0.65	−0.71	−0.70	−0.69	−0.68

Notes: Cross-correlations of trend deviations for the unemployment rate, u, and the LFP rate, l. The first four rows represent counterfactuals for u and l, and the last row represents actual values for u and l. For a counterfactual all monthly transition rates, except for the ones listed in the counterfactual column, are kept at their trend values. Statistics are calculated for quarterly averages of counterfactual monthly time series. Detrended unemployment rate and LFP rate are level deviations from trend.

lines).[17] These are the three counterfactuals for the trend deviations of the unemployment rate and LFP rate, and they approximately add up to the overall trend deviation of the two rates. In Table 3, we calculate the cross-correlations between the counterfactual unemployment and LFP rates implied by these experiments.

Past research has shown that variations in the transition probabilities between employment and unemployment are a major determinant of the unemployment rate, e.g., Shimer (2012) or Elsby, Hobijn, and Şahin (2013). This observation is confirmed by Figure 4, panel A, in that variations in these probabilities account for a substantial part of the unemployment rate variation. Figure 4, panel B, demonstrates that these variations also introduce substantial volatility into the LFP rate. In fact, the counterfactual LFP rate is more volatile than the actual LFP rate. Furthermore, variations in the transition probabilities between employment and unemployment generate a strong negative

[17] Since our trend is a symmetric moving average filter, we face a problem at the beginning and end of our sample period (see footnote 9). If for this part of the sample the deviations from a presumed trend are very large, such as is the case for the years 2007–12, then this problem is even more pronounced and our adjustment to the filter will understate deviations from trend. For this reason, we replace the calculated trend values from 2008 on with the trend values in the fourth quarter of 2007. This essentially keeps the trend unemployment rate fixed at 6.2 percent and the trend LFP rate fixed at 65.5 percent from 2008 on. Thus, our procedure is likely to overstate deviations from trend from 2008 on, especially for the LFP rate.

co-movement between the unemployment rate and the LFP rate (Table 3, first row).

The co-movement of the actual unemployment rate, with the transition probabilities between inactivity and unemployment, is such that people are more likely to join the labor force as unemployed and less likely to exit the labor force from unemployment when the unemployment rate is high. Thus, these movements simultaneously increase the unemployment rate and the LFP rate. In other words, the observed variations in transition probabilities between inactivity and unemployment contribute to the volatility of the unemployment rate, and they introduce a positive co-movement between the unemployment rate and the LFP rate (see the blue lines in Figure 4 and the second row in Table 3).

For the LFP rate, the variations of transition probabilities between employment and unemployment on the one hand, and between inactivity and unemployment on the other hand, tend to almost offset each other. This means that the joint effect of the variations in these transition probabilities is a weak positive correlation between the unemployment rate and the LFP rate (see the third row of Table 3). The stronger negative actual correlation between the unemployment rate and the LFP rate is then determined by the pattern of transition probabilities between inactivity and employment. As the unemployment rate increases, the probability of making a direct transition from inactivity to employment and vice versa declines. The effect of the reduced transition rate from inactivity tends to dominate, and the LFP rate declines. Adding this feature is enough to generate a significant negative correlation between the unemployment rate and the LFP rate (last row of Table 3).

We can interpret these results using a simplified version of the dynamics between labor market states described in the Appendix. Suppose that participants make the transition from labor market state i to labor market state j at rate $\lambda_{\iota j}$. The transition rates between employment and unemployment are λ_{EU} and λ_{UE}, and the transition rates between unemployment and inactivity are λ_{UI} and λ_{IU}. Assume also that participants can make the transition between in- and out-of-the-labor-force only by going through unemployment, that is, there are no direct transitions between employment and inactivity, $\lambda_{EI} = \lambda_{IE} = 0$.[18] For fixed transition rates, the unemployment rate and LFP rate converge

[18] In part, we can look at this as the limiting case for the observation that $\lambda_{UI} \gg \lambda_{EI}$. It is, however, also true that transitions from inactivity to employment are actually more likely than transitions from inactivity to unemployment, $\lambda_{IE} > \lambda_{IU}$.

to their steady-state values, u^* respectively l^*,

$$u^* = \frac{\lambda_{EU}}{\lambda_{EU} + \lambda_{UE}} \text{ and } l^* = \left[1 + \frac{\lambda_{UI}}{\lambda_{IU}} u\right]^{-1}.$$

In the data, monthly unemployment and LFP rates tend to be close to the steady-state values implied by their monthly transition rates.

This special case illustrates three points. First, the unemployment rate would be independent of transitions between the labor force and inactivity, if it was not for transitions between inactivity and employment. Similar to a simple two-state model of the labor market that ignores variations in the LFP rate, the unemployment rate would be determined by the transition rates between employment and unemployment. Second, even with an unemployment rate that is "exogenous" to the LFP rate, the LFP rate does depend on the unemployment rate and transition rates between unemployment and inactivity. In particular, a lower unemployment rate implies a higher LFP rate, which helps generate the observed negative correlation between the unemployment rate and the LFP rate. Third, the observed cyclical movements in the transition rates between unemployment and inactivity imply that the ratio of λ_{UI} to λ_{IU} is decreasing as the unemployment rate u increases, thereby introducing a positive correlation between the unemployment rate and the LFP rate and dampening the co-movement. Thus, transitions between employment and inactivity have to be considered if one wants to account for the co-movement between unemployment and LFP.

4. CONCLUSION

Many observers of the U.S. labor market perceive the LFP rate to be below its long-run trend and the unemployment rate to be above its long-run trend. In fact, the low cyclical LFP rate is seen as keeping the cyclical unemployment rate from being even higher, because poor employment prospects have induced discouraged unemployed workers to leave the labor force and have prevented marginally attached inactive participants from a return to the job search. In this article, we have documented that direct observations on transition rates between unemployment and out-of-the-labor-force are inconsistent with this perception. It turns out that at times of high unemployment, unemployed workers are less likely to exit the labor force and inactive workers are more likely to return to the labor force as unemployed. This pattern would have introduced a positive correlation between cyclical movements of the unemployment rate and the LFP rate. Yet we have observed a negative correlation between the two rates. We have shown that the negative co-movement is induced by movements in the unemployment rate itself, and by a procyclical transition rate from inactivity to employment without an intervening unemployment spell. To summarize, a low cyclical LFP rate to some extent simply seems to reflect a high current unemployment rate rather than to indicate an elevated future unemployment rate.

We have just described the "usual" co-movements between labor market transition rates, the unemployment rate, and the LFP rate over the business cycle. Since 2010, the unemployment rate has been declining gradually, and if we had observed the usual co-movement pattern, we should have seen the LFP rate increasing with at most a one-year lag, say, starting in 2011. We have not seen that. The LFP rate has been on a long-run declining trend since 2000, with an acceleration of that decline during the Great Recession. It is generally agreed that part of the decline in the LFP rate since 2000 reflects a demographic change that will persist over time. Current forecasts call for a further decline of the LFP rate over the next 10 years (see, for example, Toossi [2012]). But it is also argued that the more recent decline in the LFP rate reflects temporary cyclical effects that will be reversed over time (see, for example, Erceg and Levin [2013]). The recent "unusual" co-movement between the unemployment rate and LFP rate does speak to this issue. In particular, the recent observations on co-movement would appear to be less unusual if we were to attribute more of the decline in the LFP rate to a change in its long-run trend than to short-run cyclical effects.

This interpretation has implications for the medium-run forecast for gross domestic product (GDP). A falling LFP rate will dampen any increase in employment and corresponding increase in per capita GDP, even as the unemployment rate continues to decline. Thus, whereas the increasing trend for the LFP rate contributed to per capita GDP growth before 2000, the declining trend from 2000 will reduce the trend growth rate of per capita GDP. Depending how much the LFP rate is currently below trend, a return to trend might dampen this negative effect for per capita GDP growth in the near term.

APPENDIX: SOME MATH

Let $f_{ij,t}$ denote the gross flow between labor market state i in period $t-1$ and state j in period t, with $i, j \in \{E, U, I\}$. Disregarding inflows to and outflows from the working age population, the total number of people in labor market state i at time $t-1$ is

$$s_{i,t-1} = \sum_k f_{ik,t} = \sum_k f_{ki,t-2}. \tag{1}$$

The probability that a participant makes the transition from state i in period $t-1$ to state j in period t is simply

$$p_{ij,t} = f_{ij,t}/s_{i,t-1}. \tag{2}$$

The unemployment rate and LFP rate are

$$u_t = \frac{s_{U,t}}{s_{U,t} + s_{E,t}} \text{ and } l_t = \frac{s_{U,t} + s_{E,t}}{s_{U,t} + s_{E,t} + s_{I,t}}. \tag{3}$$

Conditional on initial values for the stocks, s_{i0}, we can obtain the sequence of future stocks from the sequence of transition probabilities by iterating on the equation

$$s_{i,t} = \sum_j p_{ji,t} s_{j,t-1}. \tag{4}$$

This defines a mapping from the sequence of transition probabilities, p, to the sequence of stocks, s,

$$s = G\left(p; s_0\right), \tag{5}$$

conditional on initial stocks s_0. Suppose we have a series for the trend transition probabilities, $p^T_{ij,t}$. Then we can use the above mapping to construct the implied trend values for stocks

$$s^T = G\left(p^T; s_0\right), \tag{6}$$

and we calculate the implied trend values for the unemployment rate and LFP rate, u^T and l^T.

In order to evaluate the contribution of a group of transition probabilities to the overall variation of the unemployment rate and LFP rate, we simply construct a counterfactual path for the stocks where we keep all but the probabilities of interest at their trend values and set the probabilities of interest to their actual values. For example, in order to evaluate the contribution of variations in the k-th transition probability, we construct the series

$$s^{CF}_k = G\left(p_k, p^T_{-k}; s_0\right) \tag{7}$$

with implied series for the unemployment rate and LFP rate, u^{CF}_k and l^{CF}_k. The contribution of the k-th probability to unemployment rate variations is then defined as $u^{CF}_k - u^T$.

The actual implementation of the procedure in Section 3 is slightly more complicated in that we allow for inflows and outflows to the working age population, and we replace the discrete time month-to-month transition probabilities with a continuous time process as described in Shimer (2012).

Modeling labor market transitions as a continuous time process deals with issues of time aggregation in the data. For example, if the exit rate from unemployment is relatively high, as it is most of the time, our estimates of entry probabilities to unemployment from month-to-month gross flow data might be biased since we are missing the people who do become re-employed within the month. In fact, the month-to-month transition probabilities between two particular labor market states, for example, employment and unemployment, will be an amalgam of the continuous time transition rates between all labor market states. The procedure of Shimer (2012) simply provides a way to recover the continuous time transition rates between labor market states that give rise to the observed discrete time transition probabilities.

The continuous time representation of labor market transitions also provides a convenient tool to interpret the role of transitions between unemployment and inactivity for the path of the unemployment rate and the LFP rate. The continuous time analog for the discrete time transition equation for labor market states (4) is given by

$$\begin{aligned} \dot{s}_E &= -\left(\lambda_{EU} + \lambda_{EI}\right) s_E + \lambda_{UE} s_U + \lambda_{IE} s_I \\ \dot{s}_U &= \lambda_{EU} s_E - \left(\lambda_{UE} + \lambda_{UI}\right) s_U + \lambda_{IU} s_I \\ \dot{s}_I &= \lambda_{EI} s_E + \lambda_{UI} s_U - \left(\lambda_{IE} + \lambda_{IU}\right) s_I \\ 1 &= s_E + s_U + s_I, \end{aligned} \tag{8}$$

where a dot denotes the time derivative of a variable, λ_{ij} denotes the continuous time transition rate from state i to state j, and we have normalized the size of the working age population to one. For example, on the one hand, employment declines because employed workers make the transition to unemployment at the rate λ_{EU} and exit the labor force at the rate λ_{EI}. On the other hand, employment increases because unemployed workers find employment at the rate λ_{UE} and inactive participants join the labor force and immediately find employment at the rate λ_{IE}. Subtracting outflows from inflows yields the net change of employment.

The continuous time representation of the monthly transition probabilities assumes that the transition rates remain fixed for a month. The observed transitions rates between labor market states tend to be sufficiently large such that the steady state of the system (8) for the given monthly transition rates is a good approximation of the actual stock values. The steady state of the system for fixed transition rates is an allocation of the population over labor market states such that inflows and outflows cancel and the stock values do not change, $\dot{s} = 0$. Solving equations (8) for steady-state stocks and the implied steady-state unemployment rate and LFP rate is a bit messy, but it simplifies considerably if we assume that transitions between in- and out-of-the-labor-force have to proceed through unemployment, that is, $\lambda_{EI} = \lambda_{IE} = 0$. For this case we find that the steady-state unemployment rate and LFP rate are

$$u^* = \frac{\lambda_{EU}}{\lambda_{EU} + \lambda_{UE}} \text{ and } l^* = \left[1 + \frac{\lambda_{UI}}{\lambda_{IU}} u\right]^{-1}.$$

For this special case, the unemployment rate is independent of transitions between the labor force and inactivity. Similar to a simple two-state model of the labor market that ignores variations in the LFP rate, the unemployment rate is determined by the transition rates between employment and unemployment. On the other hand, the LFP rate does depend on the unemployment rate and transition rates between unemployment and inactivity. In particular, a lower unemployment rate implies a higher LFP rate, which helps generate the observed negative correlation between the unemployment rate and the LFP rate. From Section 2 we have that the transition rates from unemployment to inactivity (inactivity to unemployment) are negatively (positively) correlated with the unemployment rate. This would imply that the LFP rate increases as the unemployment rate increases. Thus, the movements in the transition rates between in- and out-of-the-labor-force alone would yield a counterfactual positive correlation between the unemployment rate and the LFP rate.

REFERENCES

Aaronson, Stephanie, Bruce Fallick, Andrew Figura, Jonathan Pingle, and William Wascher. 2006. "The Recent Decline in the Labor Force Participation Rate and Its Implications for Potential Labor Supply." *Brookings Papers on Economic Activity* 37 (1): 69–134.

Abowd, John M., and Arnold Zellner. 1985. "Estimating Gross Labor-Force Flows." *Journal of Business & Economic Statistics* 3 (July): 254–83.

Barnichon, Regis, and Andrew Figura. 2010. "What Drives Movements in the Unemployment Rate? A Decomposition of the Beveridge Curve." Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C. Available at www.federalreserve.gov/pubs/feds/2010/201048/201048pap.pdf.

Baxter, Marianne, and Robert G. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *The Review of Economics and Statistics* 81 (November): 575–93.

Bureau of Labor Statistics. 2012. "Chapter 1. Labor Force Data Derived from the Current Population Survey." Available at http://www.bls.gov/opub/hom/pdf/homch1.pdf (September 5).

Christiano, Lawrence J., and Terry J. Fitzgerald. 2003. "The Band Pass Filter." *International Economic Review* 44 (2): 435–65.

Current Population Survey. 2000. "Design and Methodology." U.S. Census Bureau Technical Paper 63. Available at www.census.gov/prod/2000pubs/tp63.pdf.

Daly, Mary, Early Elias, Bart Hobijn, and Òscar Jordà. 2012. "Will the Jobless Rate Drop Take a Break?" *FRBSF Economic Letter* 2012-37 (December 17).

Davidson, Paul. 2013. "Drop in Labor Participation Rate is Distress Signal." *USA Today* (April 7). Available at www.usatoday.com/story/money/business/2013/04/07/march-labor-force-participation/2057887/.

Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin. 2013. "On the Importance of the Participation Margin for Labor Market Fluctuations." Federal Reserve Bank of San Francisco Working Paper 2013-05 (February).

Elsby, Michael W., Ryan Michaels, and Gary Solon. 2009. "The Ins and Outs of Cyclical Unemployment." *American Economic Journal: Macroeconomics* 1 (January): 84–110.

Erceg, Christopher J., and Andrew T. Levin. 2013. "Labor Force Participation and Monetary Policy in the Wake of the Great Recession." Available at www.bostonfed.org/employment2013/papers/Erceg_Levin_Session1.pdf.

Feng, Shuaizhang, and Yingyao Hu. 2013. "Misclassification Errors and the Underestimation of the U.S. Unemployment Rate." *American Economic Review* 103 (April): 1,054–70.

Frazis, Harley J., Edwin L. Robison, Thomas D. Evans, and Martha A. Duff. 2005. "Estimating Gross Flows Consistent with Stocks in the CPS." *Monthly Labor Review* 128 (September): 3–9.

Fujita, Shigeru, and Garey Ramey. 2009. "The Cyclicality of Separation and Job Finding Rates." *International Economic Review* 50 (2): 415–30.

Hatzius, Jan. 2012. "Still Lots of Slack in the Labor Market." *US Economics Analyst* 12/41. New York: Goldman Sachs Economics Research (October 12).

Hodrick, Robert J., and Edward C. Prescott. 1997. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit, and Banking* 29 (February): 1–16.

Poterba, James M., and Lawrence H. Summers. 1986. "Reporting Errors and Labor Market Dynamics." *Econometrica* 54 (November): 1,319–38.

Shimer, Robert. 2012. "Reassessing the Ins and Outs of Unemployment." *Review of Economic Dynamics* 15 (April): 127–48.

Tankersley, Jim. 2013. "Vanishing Workforce Weighs on Growth." *Washington Post* (April 6). Available at www.washingtonpost.com/business/economy/vanishing-workforce-weighs-on-growth/2013/04/06/2bc46116-9e20-11e2-9a79-eb5280c81c63_story.html?tid=pm_business_pop.

Toossi, Mitra. 2012. "Labor Force Projections to 2020: A More Slowly Growing Workforce." *Monthly Labor Review* 135 (January): 43–64.

Economic Quarterly—Volume 99, Number 1—First Quarter 2013—Pages 25–43

A Cohort Model of Labor Force Participation

Marianna Kudlyak

The aggregate labor force participation (LFP) rate measures the share of the civilian noninstitutionalized population who are either employed or unemployed (i.e., actively searching for work). From 1963 to 2000, the LFP rate was rising, reaching its peak at 67.1 percent. The LFP rate has been declining ever since, with the decline accelerating after 2007. Between December 2007 and December 2012, the LFP rate declined from 66 percent to 63.6 percent. Prior to 2012, the last year when the LFP rate was below 65 percent was 1986.

The decline in the LFP rate, which coincided with the Great Recession, raises the question: Is the LFP rate at the end of 2012 close to or below its long-run trend? The question is important to policymakers and economists. If a large portion of the workers who are currently out of the labor force represents workers who are temporarily out of the labor force, then the unemployment rate by itself might not be a good measure of the slack in the economy.

In this article, we discuss the change in the aggregate LFP rate from 2000 to 2012, with an emphasis on the changes in the age-gender composition of the population and changes in the LFP rates of different demographic groups. We then estimate a cohort-based model of the LFP rates of different age-gender groups and construct the aggregate LFP rate using the model estimates. The model is a parsimonious version of the model studied in Aaronson et al. (2006). It contains age-gender effects, birth-year cohort effects, and the estimated deviations of employment from its long-run trend as the cyclical indicator. We

The author is grateful, without implicating them in any way, to Bob Hetzel, Andreas Hornstein, Marios Karabarbounis, Steven Sabol, and Alex Wolman for their comments. The author thanks Peter Debbaut and Samuel Marshall for excellent research assistance. The views expressed here are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: marianna.kudlyak@rich.frb.org.

estimate the model on the 1976–2007 data and then predict the aggregate LFP rate for 2008–12.

We find that in 2008–11, the actual LFP rate closely follows the LFP rate predicted from the model that takes into account the estimated cyclical deviation of employment from its trend. In 2012, the actual LFP rate is in fact above the estimated value from the model. The actual LFP rate in 2012 is close to the estimated trend constructed from the actual age-gender composition of the population and the age-gender and cohort effects estimated from the model.

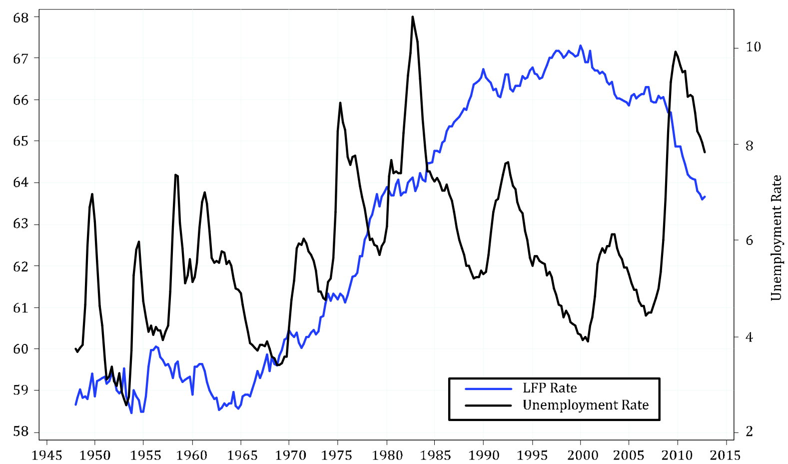
What are the factors behind the LFP rate in 2012 being above the value predicted from the model with the cyclical indicator? In the model, we use estimated deviations of employment from its long-run trend as a cyclical indicator. While it is true that the decline in employment during the Great Recession contributed to lowering labor force participation in 2008–12, it also appears that other factors during the 2007–09 recession worked to counteract this effect in 2012. Our model is silent about these factors. One can speculate that the increase in the duration of unemployment insurance benefits, or the decline in household wealth (due to the collapse of stock and housing markets), might have contributed to workers remaining in the labor force at a larger rate than predicted by the cyclical component of employment.

This article is related to an active debate in the recent academic and policy circles. The theoretical models are studied in Veracierto (2008), Krusell et al. (2012), and Shimer (2013). The empirical discussions are provided in Kudlyak, Lubik, and Tompkins (2011); Aaronson, Davis, and Hu (2012); Daly et al. (2012); Hotchkiss, Pitts, and Rios-Avila (2012); Canon, Kudlyak, and Debbaut (2013); and Schweitzer and Tasci (2013). The cohort model employed in the modeling labor force participation rate was originally proposed by Aaronson et al. (2006). Fallick and Pingle (2006) and Balleer, Gómez-Salvador, and Turunen (2009) provide extensions to the model.

The findings in the article are consistent with the findings in Aaronson et al. (2006), whose 2006 projection of the LFP rate in 2012 is 63.7 percent, the number that coincides with the actual rate in 2012. Other studies find that the LFP rate in 2012 is below its trend (Aaronson, Davis, and Hu [2012]; Bengali, Daly, and Valletta [2013]; Erceg and Levin [2013]; Hotchkiss and Rios-Avila [2013]).

The rest of the article is structured as follows. The first section reviews the behavior of the aggregate LFP rate during 2000–12 and presents counterfactual exercises using an age-gender decomposition of the aggregate LFP rate. Section 2 describes the cohort model and presents the empirical results. Section 3 concludes.

Figure 1 LFP Rate and Unemployment Rate



Notes: Quarterly averages of seasonally adjusted (SA) monthly series, January 1949–December 2012. Author's calculations using series from HAVER.

1. WHAT COMPONENTS DRIVE THE CHANGES IN THE AGGREGATE LFP RATE DURING 2000–12?

After reaching its peak of 67.3 percent in the first half of 2000, the aggregate LFP rate declined from 2000 to Q2:2004, stabilized for a few years, and then started falling again in 2008.[1] Figure 1 shows the aggregate LFP rate and the aggregate unemployment rate.

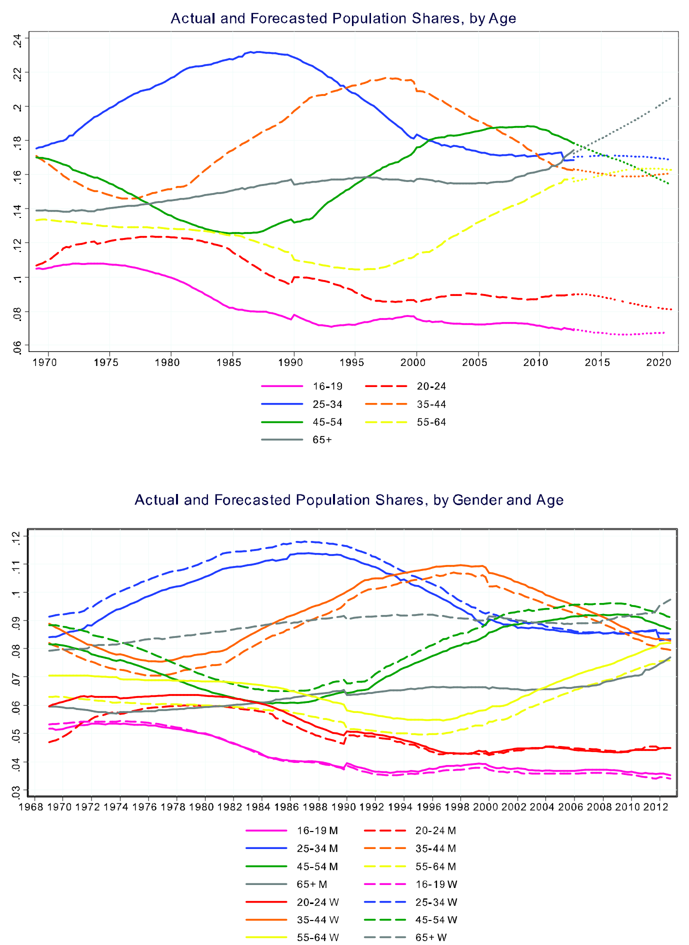
The aggregate LFP rate can be decomposed into the weighted sum of the LFP rates of different demographic groups, i.e.,

$$LFP_t = \sum_i s_t^i LFP_t^i, \tag{1}$$

where LFP_t^i is the labor force participation rate of group i, s_t^i is the population share of group i, i.e., $s_t^i \equiv \frac{Pop_t^i}{Pop_t}$, and Pop_t^i is the population of group i.

[1] The data reported in the article are from HAVER (SA), unless stated otherwise. The last data point at the time of the analysis: December 2012.

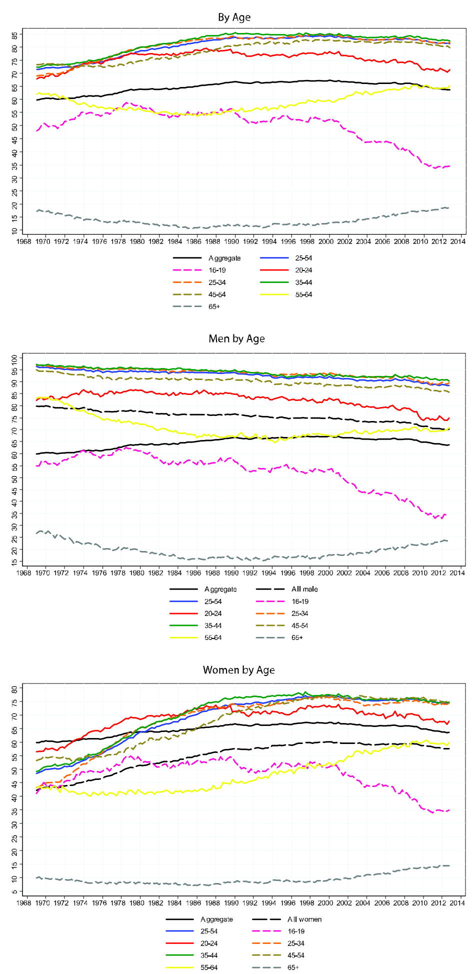
Figure 2 Actual and Forecasted Population Shares



Notes: Quarterly averages of monthly series, January 1969–December 2012. Annual series for estimates. All series from HAVER. Dotted lines denote shares based on the HAVER-estimated forecast of resident; population adjusted by the author to represent civilian noninstitutionalized population (using 2012).

To understand what forces drove the decline of the LFP rate since 2008, we first examine the change in the demographic composition of the population and the change in the LFP rates of different age-gender groups. Figure 2 shows the population shares by age-gender group.

Figure 3 LFP Rates



Notes: Quarterly averages of monthly series, January 1969–December 2012. All series from HAVER. Dotted lines denote shares based on the HAVER-estimated forecast of resident; population adjusted by the author to represent civilian noninstitutionalized population (using 2012). Author's calculations from the population and labor force series, SA.

Figure 3 shows the LFP rates of different age-gender groups. As can be seen from the figures, the developments that took place between Q4:2007 and Q4:2012 are a continuation of the developments that have been taking place since 2000, when the aggregate LFP rate reached its peak:[2]

- The composition of the population has been shifting toward older workers who typically have lower labor force attachment. This is in part due to the population of baby boomers gradually moving from the prime working age group with a high LFP rate to older age groups with lower LFP rates. Also note that the share of older women is larger than the share of older men, and women typically have lower labor force attachment than men.
- The LFP rate of 25- to 54-year-old workers, a group with the highest LFP rate, has been declining. From Q4:2007 to Q4:2012, the rate declined from 82.9 percent to 81.3 percent.
- The LFP rate of teenagers and young adults has been declining.
- The LFP rate of women has started to decline after increasing prior to 1999.
- The LFP rate of men has continued its decline, which started in the 1940s.

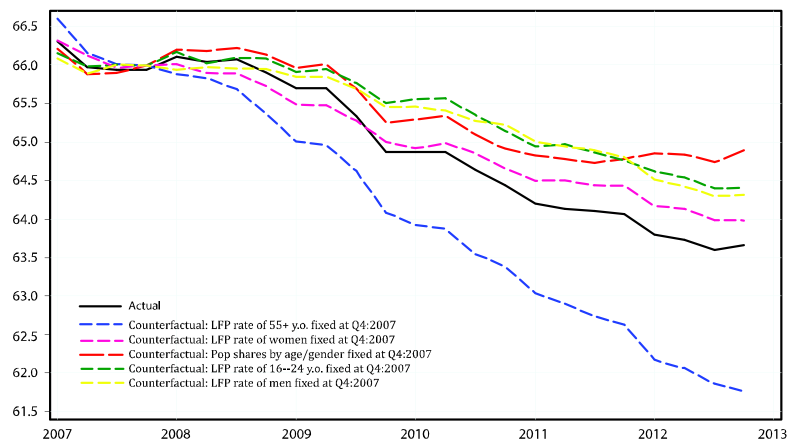
How Much Change Is Driven by the LFP Rates of Different Demographic Groups?

To understand the importance of the compositional changes and of the changes in the labor force participation rates of different demographic groups, we first present counterfactual exercises to quantify the impact of these changes on the aggregate labor force participation rate.

In the exercises, we keep the LFP rate of specific demographic groups fixed at their Q4:2007 level and allow the LFP rates of all other groups and the demographic composition of the population to follow their actual path. We consider four such counterfactual exercises: (1) fixing the LFP rate of 55+ year-old workers, (2) fixing the LFP rate of 16- to 24-year-old workers, (3) fixing the LFP rate of women, and (4) fixing the LFP rate of men. These exercises demonstrate the importance of changes in the LFP rates of different demographic groups for changes in the aggregate LFP rate. In our fifth counterfactual exercise, we fix the population shares of age-demographic groups at their Q4:2007 levels and allow the groups' LFP rates to follow their actual path. The results of these exercises are shown in Figure 4.

[2] See Toossi (2012a, 2012b) for a description of the trends and Canon, Kudlyak, and Debbaut (2013) for a summary of the Bureau of Labor Statistics projections.

Figure 4 Labor Force Participation Rate, Actual and Counterfactual



Notes: Author's calculations using data from HAVER, quarterly averages of monthly series.

As can be seen from the figure, the experiment with holding the LFP rates of 55–64 and 65+ year-old workers fixed (the dashed blue line) delivers the largest discrepancy between the actual aggregate LFP (the solid black line) and the counterfactual one. Since the LFP rate of older workers has increased, the counterfactual rate lies below the actual LFP rate, and in Q4:2012 stands at 61.7 percent.

The second largest discrepancy (in absolute value) between the actual aggregate LFP and the counterfactual one is obtained from holding the population shares fixed at their 2007 levels (the dashed red line). In this case, the counterfactual LFP rate exceeds the actual one and stands at almost 65 percent in Q4:2012. We see that between 2007 and 2012 the population composition has shifted toward a composition with lower labor force attachment.

The results also show that the counterfactual based on the fixed LFP of 16- to 24-year-old workers (the dashed green line) and the counterfactual based on the fixed LFP of men (the yellow dashed line)

line up almost perfectly and are both above the actual aggregate LFP rate.

Finally, the figure shows that the counterfactual LFP rate based on the fixed LFP by women (the dashed pink line) has declined more than the one based on the fixed LFP by men (the dashed yellow line), while both counterfactuals lie above the actual LFP rate.

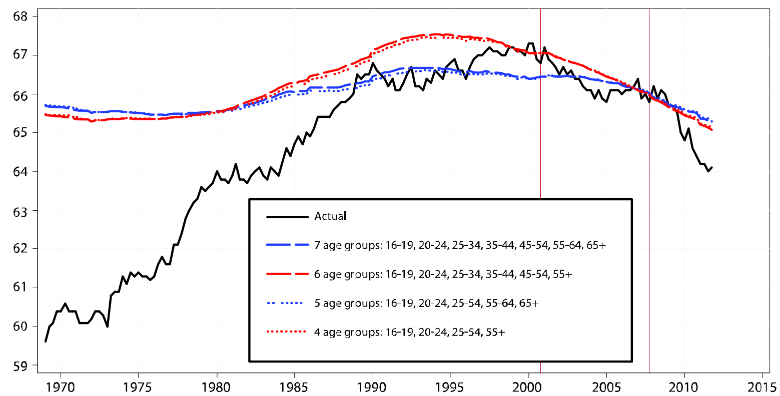
An Alternative Decomposition of LFP

As an alternative way of gauging how much of the change in the LFP rate was driven by the change in the population shares of different demographic groups, we perform the following counterfactual. We fix the LFP rates of 14 age-gender groups at their respective levels at time t_0 and construct the counterfactual LFP rate using the actual population shares of the respective groups, i.e., $LFP_t^{t_0} = \sum_i s_t^i LFP_{t_0}^i$. In the analysis, we consider the following seven age groups for each gender: 16–19, 20–24, 25–34, 35–44, 45–54, 55–64, and 65 and older. The blue lines in Figures 5 and 6 show the counterfactual LFP for t_0 equal to Q4:2007 and t_0 equal to Q4:2000, respectively.

As can be seen from Figure 6, in Q4:2012, the counterfactual LFP rate constructed from the groups' LFP rates fixed at their levels in Q4:2000 is 65.5 percent, while from 2000 to 2012 the actual LFP rate declined from 67 percent to 63.6 percent. The counterfactual LFP rate constructed from the age-gender LFP rates fixed at their levels in Q4:2007 is 65 percent, while from 2007 to 2012 the actual LFP rate declined from 66 percent to 63.6 percent (Figure 5). Thus, the results suggest that the demographic change of the population is associated with approximately 40 percent of the decline of the aggregate LFP rate between 2000 and 2012 and 37 percent of the decline between 2007 and 2012.

For such demographic counterfactuals it is important to consider as fine a group classification as possible, especially if there are substantial differences in the LFP rates of workers of different ages combined into a group. For example, the red lines in Figures 5 and 6 show the counterfactual LFP rate when we consider only six age groups for each gender (16–19, 20–24, 25–34, 35–44, 45–54, and 55 and older), i.e., combining ages 55–64 and 65+ into one group, 55+. As can be seen from the figures, in this case LFP_t^{2000} has declined more than the counterfactual rate in the seven-age-group exercise (64.4 percent). This is because the share of 55- to 64-year-old workers in the 55+ group, who have much higher labor force attachment than 65+, has increased between 2000 and 2012 (see Figure 6 for the shares).

Figure 5 The Counterfactual LFP Rate based on the Change in the Demographic Composition of the Population, Q4:2007



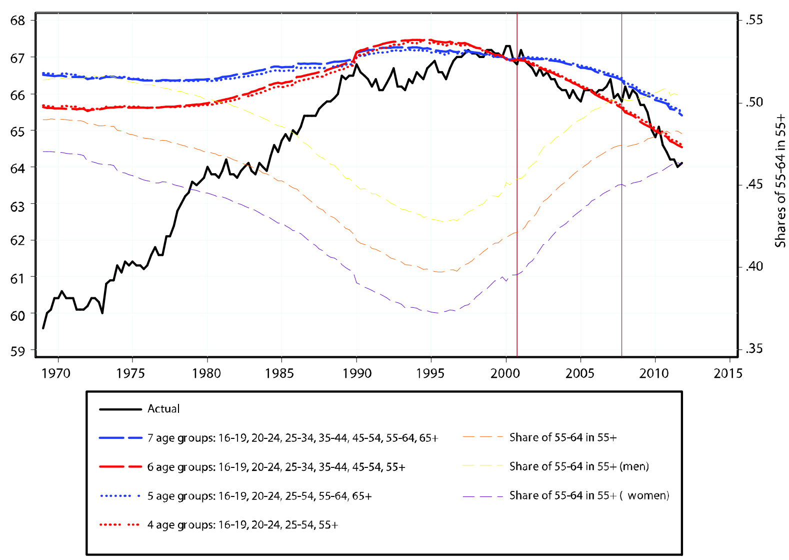
Notes: Population forecast is based on residential population forecast from HAVER, scaled by 2012 relationship between residential and civilian noninstitutionalized population by age and gender.

The observations above show that the demographic composition of the population and the changes in the LFP rates of different groups have played an important role in the change of the aggregate LFP rate. We now proceed to examine the age-gender and cohort effects in the LFP rates of different demographic groups on the aggregate LFP rate.

2. A COHORT-BASED MODEL OF LABOR FORCE PARTICIPATION

The results in Section 1 show that the time-variation in the LFP rates of different demographic groups are important for the variation in the aggregate LFP rate. In this section, we propose a model for the trend in the LFP rates of different demographic groups. We then estimate the trend in the aggregate LFP rate using the estimated trends in the

Figure 6 The Counterfactual LFP Rate based on the Change in the Demographic Composition of the Population, Q4:2000



Notes: Population forecast is based on residential population forecast from HAVER, scaled by 2012 relationship between residential and civilian noninstitutionalized population by age and gender.

LFP rates of different demographic groups and the actual demographic composition of the population.

Model

Life-Cycle and Cohort Effects in the LFP Rates of Age-Gender Groups

The LFP rates of different demographic groups reflect life-cycle and gender effects. In addition to these effects, the year-of-birth cohort effects can be an important determinant of the labor force attachment of a demographic group in a particular period. For example, as noted earlier, the baby boomers typically have higher labor force attachment.

As this cohort ages and moves through the age distribution, its stronger labor force attachment carries over to the respective age group.

We think of the demographic and the cohort effects in the LFP rates of different demographic groups as the determinants of the long-run labor force participation trend. To estimate this trend, we specify the following model:

$$\ln LFP_t^i = \alpha + \ln \alpha_i + \frac{1}{n} \sum_{b=1917}^{1996} C_{b,i,t}^f \ln \beta_b^f + \frac{1}{n} \sum_{b=1917}^{1996} C_{b,i,t}^m \ln \beta_b^m + \varepsilon_{i,t}, \quad (2)$$

where LFP_t^i is the labor force participation rate of age-gender group i, α_i is the fixed effect of age-gender group i, $C_{b,i,t}^f$ is the dummy variable that takes value 1 if age-gender group i in period t includes women born in year b, $C_{b,i,t}^m$ is the dummy variable that takes value 1 if age-gender group i in period t includes men born in year b, and n denotes the number of ages in group i. We specify separate cohort effects for men and women, i.e., $\beta_b^f (\beta_b^m)$ is the cohort-specific fixed effect of a cohort of women (men) born in year b. We assume that each cohort has equal importance in the corresponding age group conditional on the number of cohorts in the group. For the oldest group, 65+, we set $n = 20$ (setting $n = 30$ does not have a substantial effect on the results). To identify age-gender and cohort effects, we normalize $\ln \alpha_1 = 0$ and $\ln \beta_{1969}^f = 0$. The model is estimated using pooled quarterly data on the LFP rates of 14 age-gender groups.

The model in equation (2) is a simplified version of a model in Aaronson et al. (2006). Using the estimates from equation (2), we obtain the time series of $\ln LFP_t^i$ for the 14 age-gender groups, $\widehat{\ln LFP_t^i}$, and calculate $\widehat{LFP_t^i} = \exp\left(\widehat{\ln LFP_t^i} + \frac{\sigma_\varepsilon^2}{2}\right)$, where σ_ε^2 is the variance of $\widehat{\varepsilon_{i,t}}$. We then construct the estimated aggregate LFP rate as

$$\widehat{LFP_t} = \sum_i s_t^i \widehat{LFP_t^i}, \quad (3)$$

where s_t^i denotes the actual population share of group i in quarter t. Thus, the population shares capture the effect of the change in the demographic composition of the labor force, while $\widehat{LFP_t^i}$ reflects the age-gender and cohort effects of the different demographic groups. We refer to $\widehat{LFP_t}$ from model (2) as the estimated trend in the aggregate LFP rate.

Life-Cycle, Cohort, and Cyclical Effects

To further understand the behavior of the aggregate LFP rate, we also estimate a model similar to the one in equation (2) with a cyclical indicator. The cyclical indicator is the percentage deviation of employment from its trend. The idea behind the indicator is that when the labor market is weak, the labor force participation declines.[3]

The cohort model with the cyclical indicator is

$$\ln LFP_t^i = \alpha + \ln \alpha_i + \frac{1}{n}\sum_{b=1917}^{1996} C_{b,i,t}^f \ln \beta_b^f + \frac{1}{n}\sum_{b=1917}^{1996} C_{b,i,t}^m \ln \beta_b^m + \sum_{g=1}^{14} I(i=g)\left(d\ln E_t \ln \gamma_g^0 + d\ln E_{t-1}\ln \gamma_g^1 + d\ln E_{t-2}\ln \gamma_g^2\right) + \varepsilon_{i,t}, \tag{4}$$

where $I(\cdot)$ is the indicator function, and $d\ln E_t$ is the percentage deviation of the employment series from its Hodrick-Prescott (HP)-filtered trend with a smoothing parameter $\lambda = 10^5$ applied to the quarterly data.

In the estimation, we use the contemporaneous percentage deviation from employment as well as the first and second lag of the deviation. Note that we allow the cyclical effects to vary by demographic group i. Because of the end-of-sample issues associated with HP-filtering the series, we experiment with using a counterfactual cyclical series, $\widetilde{d\ln E_t}$, obtained by calculating the deviations from the employment series simulated to grow at the 2 percent year-over-year quarterly rate after Q4:2012. While the cyclical components from the actual and simulated employment series differ after 2009, the model-based aggregate LFP rates from the two alternative series are very similar.

The model is estimated on quarterly data. After estimating equation (4), we construct the aggregate LFP rate as described in equation (3).

The error term in equation (4), $\varepsilon_{i,t}$, captures the residual between the actual LFP rate of group i in period t and the one explained by the historical relationship between age-gender, cohort, and cyclical effects and the LFP rates by group. Thus, the residual captures two main effects. First, it captures the factors that affect the LFP of group i that are not modeled explicitly in equation (4). These include some structural factors (for example, changes in taxes or disability benefits) and some cyclical factors that are not fully captured by the changes in aggregate employment (for example, changes in the duration of unemployment benefits, house prices, and stock prices). Second, the residual captures potential changes in individuals' behavior (i.e., changes in responses of the LFP rates to different structural and cyclical factors).

[3] See recent evidence in Hotchkiss, Pitts, and Rios-Avila (2012); Kudlyak and Schwartzman (2012); Elsby, Hobijn, and Şahin (2013); and Hornstein (2013).

Empirical Results

One way to obtain the predictions from the models described in equations (2) and (4) is to estimate the models using the 1976–2012 data, obtain the trend in the aggregate LFP rate (from equation [3]) and the model-predicted aggregate LFP rate from the model with a cyclical indicator, and compare the estimates with the actual LFP rate during 2008–12. Another way is to estimate the model on the 1976–2007 data and then use the estimates together with the assumptions on cohort effects and predict the aggregate LFP rate for 2008–12. The cohort model is sensitive to which approach is used.

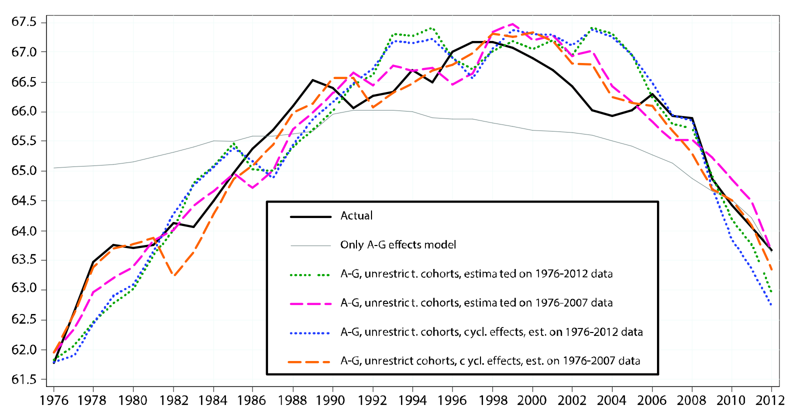
One of the concerns associated with cohort models is the end-of-the-sample effect. In particular, the young cohorts observed in the 1976–2012 sample (i.e., those born in 1985–1996) are observed only during the period of the declining aggregate LFP rate. Thus, the model identifies these cohorts' propensity to participate from the period of overall low participation, attributing low LFP to these young cohorts rather than to the model's residual. Given the severity and the length of the Great Recession, the effects of the cohorts born prior to 1985 are also, to a large extent, identified from their labor force participation rates during 2008–2012, the period of the overall low LFP. This is the case for cohorts for which, for example, at least half of the observations come from the 2008–12 period.

To avoid the end-of-sample effect on the estimates, we estimate the models in equations (2) and (4) using the data from 1976–2007. To construct the prediction of the aggregate LFP rate for 2008–12, we assign, for cohorts born after 1991, the average cohort effect of the last 20 cohorts. Figure 7 shows the following series: (1) the actual aggregate LFP rate, (2) the LFP rate constructed from the model with only age-gender effects, (3) the LFP rate constructed from the model with age-gender and cohort effects estimated on 1976–2007 data, and (4) the LFP rate constructed from the model with age-gender, cohort, and cyclical effects estimated on 1976–2007 data.[4]

As can be seen from the figure, the aggregate LFP rate estimated from the model with only age-gender and cohort effects on the 1976–2007 sample exceeds the actual aggregate LFP rate after 2008, and the two lines coincide at the end of 2012. This measure constitutes our preferred measure of the trend in the LFP rate. The aggregate LFP rate estimated from the model with age-gender, cohort, and cyclical

[4] The estimates are available from the author.

Figure 7 Actual and Model-Based Aggregate LFP Rate, Age-Gender and Cohorts Effects



Notes: To construct the LFP from the model estimated on the 1976–2007 data, we estimate unrestricted cohort effects for birth years from 1917 to 1991 and then assign the average cohort effect of the last 20 cohorts to cohorts born in 1992–96.

effects on the 1976–2007 data closely tracks the actual aggregate LFP rate during 2008–11 and is slightly below it in the last quarter of 2012.

For comparison, Figure 7 also shows the aggregate LFP rate estimated from the models using the 1976–2012 data. As can be seen from the figure, during 2008–12, the aggregate LFP rate predicted from the model estimated using the 1976–2007 data exceeds the aggregate LFP rate predicted from the model estimated using the 1976–2012 data. This is true for the predictions from the model with age-gender and cohort effects and for the predictions from the model with age-gender, cohort, and cyclical effects. It appears that the model estimated using the 1976–2012 data attributes the cyclical effects of the 2008–12 period to cohort effects. To minimize the end-of-sample effect, we also estimated the models employing a restriction on cohorts as described in Aaronson et al. (2006). In particular, we constrain the evolution of the fixed effects for consecutive pairs of the cohorts born in 1985–96 so that the difference in the average propensity to participate between one cohort and the next is the same as for a set of cohorts observed over the last full business cycle. The aggregate LFP rate based on the models with restricted and unrestricted cohorts are similar, so the figure shows only the results without restrictions.[5]

Discussion

In the model, the cohort effect stands for an average effect of all non-modeled factors (beyond life-cycle, gender, and cyclical effects) that affect the labor force participation of a cohort (i.e., the workers born in a particular year) throughout the period the cohort is observed in the sample. These factors can include both structural and cyclical variables. For example, the availability of and the rules that govern Social Security benefits and disability insurance might influence the decision to look for work versus drop out of the labor force. The wage premium from higher educational attainment might influence the decision of younger workers to go to school rather than participate in the labor force. The availability and cost of child care can influence the decision of mothers to join the labor force.

Consequently, the cohort effects constitute a black box that aggregates these influences and serve as a useful device for accounting exercises. The cohort model, however, might not be the best laboratory for long-term forecasts. In our estimation, we recognize explicitly that the effect of young cohorts is to a large degree identified from the few years during which we observe these cohorts in the data. In particular, for the youngest cohorts, a low cohort effect can be due to the true low propensity of these cohorts to participate or due to the model attributing low cyclical LFP to the cohort effect. In our exercise, we control for these effects. A forecasting exercise would inevitably involve assumptions about the cohort effects going forward. It is possible that, for example, the youngest cohorts who are not participating currently due to schooling will, in fact, increase their LFP as they grow older. The cohort model does not provide information to support or reject such scenarios.

[5] The result with restrictions is available from the author. This result motivates estimation of the benchmark model (i.e., using the 1976–2007 data) without restrictions on cohorts.

3. CONCLUSION

We find that in the aftermath of the Great Recession, the aggregate LFP rate closely tracks the one predicted by the historical relationship between the changes in employment and the labor force participation rates of different age-gender groups in a cohort-based model. In 2012, the actual LFP rate is slightly higher than the one predicted by the model. In 2009–11, the trend component of the labor force participation rate, which is based entirely on the life-cycle and cohort effects of the LFP rates of different age-gender groups and the actual age-gender composition of the population, exceeds the actual LFP rate.

The result that the LFP rate in 2012 is above the level that is predicted by the historical relationship between labor force participation and the cyclical indicator is consistent with the recent findings by Hotchkiss and Rios-Avila (2013), who provide direct evidence that some changes in behavior took place.[6] What other factors could have contributed to the estimated deviation of the actual LFP rate from its model-based prediction? We speculate that the Great Recession was characterized by unusually wild swings in some economic indicators that could have affected labor force participation. First, the unemployment benefits in some states were extended to unusually high levels. The benefits extension might have kept some workers in the labor force for up to two years to enable them to collect benefits rather than dropping out of the labor force. In particular, Farber and Valletta (2013) find that the effect of the unemployment insurance extensions on unemployment exits and duration is primarily due to a reduction in exits from the labor force.[7] Second, the collapse of the stock market led to a decline in retirement savings, which might have led older workers to stay in the labor force longer. Third, the collapse of the housing market lowered the ability of households to borrow against their home equity, which also might have caused individuals to join and/or remain in the labor force at higher rates than historically predicted by age, gender, cohort, and cyclical employment effects. Finally, to understand the behavior of labor force participation and its trend, more research is needed that would explicitly model and account for the factors that

[6] In particular, Hotchkiss and Rios-Avila (2013) use microdata from the Current Population Survey and estimate the probability of an individual participating in the labor force as a function of age, education, and other socioeconomic and demographic characteristics of the individual as well the aggregate labor market conditions. They find that the coefficients on the socioeconomic and demographic characteristics estimated from the post-2008–09 period differ from the coefficients estimated from the pre-recession period in such a way as to increase the aggregate LFP rate.

[7] See also Fujita (2010, 2011) and Rothstein (2011).

influence the labor force participation decision of different demographic groups.

REFERENCES

Aaronson, Daniel, Jonathan Davis, and Luojia Hu. 2012. "Explaining the Decline in the U.S. Labor Force Participation Rate." Federal Reserve Bank of Chicago *Chicago Fed Letter* no. 296 (March).

Aaronson, Stephanie, Bruce Fallick, Andrew Figura, Jonathan Pingle, and William Wascher. 2006. "The Recent Decline in the Labor Force Participation Rate and Its Implications for Potential Labor Supply." *Brookings Papers on Economic Activity* 37 (Spring): 69–134.

Balleer, Almut, Ramón Gómez-Salvador, and Jarkko Turunen. 2009. "Labour Force Participation in the Euro Area: A Cohort Based Analysis." European Central Bank Working Paper 1049 (May).

Bengali, Leila, Mary Daly, and Rob Valletta. 2013. "Will Labor Force Participation Bounce Back?" Federal Reserve Bank of San Francisco *Economic Letter* 2013–14 (May).

Canon, Maria E., Marianna Kudlyak, and Peter Debbaut. 2013. "A Closer Look at the Decline in the Labor Force Participation Rate." Federal Reserve Bank of St. Louis *The Regional Economist* (October).

Daly, Mary, Early Elias, Bart Hobijn, and Òscar Jordà. 2012. "Will the Jobless Rate Drop Take a Break?" Federal Reserve Bank of San Francisco *Economic Letter* 2012-37 (December).

Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin. 2013. "On the Importance of the Participation Margin for Labor Market Fluctuations." Federal Reserve Bank of San Francisco Working Paper 2013-05 (February).

Erceg, Christopher J., and Andrew T. Levin. 2013. "Labor Force Participation and Monetary Policy in the Wake of the Great Recession." Mimeo.

Fallick, Bruce, and Jonathan Pingle. 2006. "A Cohort-Based Model of Labor Force Participation." Federal Reserve Board of Governors Finance and Economics Discussion Series 2007-09 (December).

Farber, Henry S., and Robert G. Valletta. 2013. "Do Extended Unemployment Benefits Lengthen Unemployment Spells? Evidence from Recent Cycles in the U.S. Labor Market." Federal Reserve Bank of San Francisco Working Paper 13-09 (April).

Fujita, Shigeru. 2010. "Economic Effects of the Unemployment Insurance Benefit." Federal Reserve Bank of Philadelphia *Business Review* Q4:20–7.

Fujita, Shigeru. 2011. "Effects of Extended Unemployment Insurance Benefits: Evidence from the Monthly CPS." Federal Reserve Bank of Philadelphia Working Paper 10-35/R (January).

Hornstein, Andreas. 2013. "The Cyclicality of the Labor Force Participation Rate." Mimeo.

Hotchkiss, Julie L., M. Melinda Pitts, and Fernando Rios-Avila. 2012. "A Closer Look at Nonparticipants During and After the Great Recession." Federal Reserve Bank of Atlanta Working Paper 2012-10 (August).

Hotchkiss, Julie L., and Fernando Rios-Avila. 2013. "Identifying Factors Behind the Decline in the U.S. Labor Force Participation Rate." *Business and Economic Research* 3 (March): 257–75.

Krusell, Per, Toshihiko Mukoyama, Richard Rogerson, and Ayşegül Şahin. 2012. "Gross Worker Flows over the Business Cycle." Mimeo (December).

Kudlyak, Marianna, Thomas Lubik, and Jonathan Tompkins. 2011. "Accounting for the Non-Employment of U.S. Men, 1968–2010." Federal Reserve Bank of Richmond *Economic Quarterly* 97 (4):359–87.

Kudlyak, Marianna, and Felipe Schwartzman. 2012. "Accounting for Unemployment in the Great Recession: Nonparticipation Matters." Federal Reserve Bank of Richmond Working Paper 12-04 (June).

Rothstein, Jesse. 2011. "Unemployment Insurance and Job Search in the Great Recession." *Brookings Papers on Economic Activity* 43 (Fall): 143-213.

Schweitzer, Mark E., and Murat Tasci. 2013. "What Constitutes Substantial Employment Gains in Today's Labor Market?" Federal Reserve Bank of Cleveland *Economic Commentary* 2013-07 (June 7).

Shimer, Robert. 2013. "Job Search, Labor Force Participation, and Wage Rigidities." In *Advances in Economics and Econometrics: Theory and Applications*, Tenth World Congress, Vol. II, Chapter 5, edited by Daron Acemoglu, Manuel Arellano, and Eddie Dekel. New York: Cambridge University Press, 197–235.

Toossi, Mitra. 2012a. "Labor Force Projections to 2020: A More Slowly Growing Workforce." *Monthly Labor Review* 135 (January): 43–64.

Toossi, Mitra. 2012b. "Projections of the Labor Force to 2050: A Visual Essay." *Monthly Labor Review* 135 (October): 3–16.

Veracierto, Marcelo. 2008. "On the Cyclical Behavior of Employment, Unemployment and Labor Force Participation." *Journal of Monetary Economics* 55 (September): 1,143–57.

Economic Quarterly—Volume 99, Number 1—First Quarter 2013—Pages 45–81

Saving for Retirement with Job Loss Risk

Borys Grochulski and Yuzhe Zhang

In this article, we study optimal saving and consumption decisions. The optimal saving problem is among the most basic questions in economics and finance. How does one best decide on what portion of income they should consume now and what portion they should save for their future consumption needs? One important aspect of this question concerns saving for retirement. What is an optimal plan for saving enough to be able to retire? In particular, how does this plan depend on the risk of losing one's job? How much more should one save if the risk of becoming jobless increases?

Our primary objective in this article is to review several important results from the general theory of optimal consumption and saving decisions, as well as provide some novel analysis of the problem of saving for retirement in particular. The problem of optimal timing of retirement is most conveniently studied in a continuous-time framework, which we employ for our analysis. Our secondary objective is to provide an accessible exposition of the techniques useful in solving continuous-time models of the type we examine.

The basic framework economists have used to study the intertemporal tradeoff between current and future consumption has the following structure. An economic agent earns a stream of labor income that can change stochastically over time. At each point in time, the agent allocates his labor income to either current consumption or to savings. The agent's preferences over consumption streams are represented by a concave utility function, i.e., the agent is averse to fluctuations in his

Grochulski is a senior economist with the Richmond Fed. Zhang is an assistant professor with Texas A&M University. The authors would like to thank Arantxa Jarque, Marianna Kudlyak, Andrew Owens, and Ned Prescott for their helpful comments. The views expressed in this article are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: borys.grochulski@rich.frb.org; yuzhe-zhang@econmail.tamu.edu.

consumption. The duration of the agent's career is in the basic model approximated by infinity, i.e., the agent earns labor income and consumes indefinitely into the future. The portion of his labor income that the agent does not immediately consume adds to his financial wealth. In the basic model, there is only one asset in which all of the agent's financial wealth is invested. The asset pays off a riskless rate of return equal to the agent's intertemporal rate of time preference—the rate of return with which the agent's optimal consumption path absent all uncertainty would be constant forever.

The model we study in this article extends this basic framework by adding to the optimal consumption and saving decision a labor supply decision operating on the extensive margin, meaning we allow the agent to stop working. If the agent quits, he loses his labor income but gains leisure. The decision to quit is irreversible, so quitting means retiring. In retirement, the agent lives off of his savings and enjoys leisure. As in the basic model, the agent remains infinitely lived in our analysis.

For tractability and ease of exposition, we assume in our model a particularly simple stochastic structure for the agent's labor income process. The agent earns a constant stream of labor income for as long as he is not fired. If he is fired, he earns nothing and cannot go back to working ever again. Thus, being fired is in our model equivalent to being sent to involuntary retirement. The observed time path of the agent's labor income in our model is thus constant, at some positive level, until the agent either is fired or quits. Afterward, it is also constant at the level of zero.

Ljungqvist and Sargent (2004, Ch. 16) review the solution to the optimal consumption and saving problem in the basic framework with income fluctuating stochastically but without retirement. The main property of the optimal consumption plan is unbounded growth of financial wealth and consumption: Provided that the labor income process does not settle down in the long run (rather, it remains sufficiently stochastic), in almost all possible resolutions of uncertainty, the amount of financial wealth the agent holds and the amount the agent consumes grow over time without bound. When we allow for endogenous retirement, this property of the optimal wealth accumulation and consumption plan no longer holds. In all possible resolutions of uncertainty, wealth and consumption converge to a finite limit.

The intuition behind this result is simple. We show that the agent's optimal retirement plan takes the form of a wealth threshold rule: The agent retires as soon his accumulated financial wealth reaches a certain threshold. With this rule, wealth will not grow without bound prior to retirement. With finite wealth and no labor income after retirement, the agent's optimal consumption also remains bounded in the long run.

In fact, consumption is constant and equal to the amount of interest income generated by the agent's wealth in retirement.

The dynamics of consumption and savings are in our model as follows. Wealth and consumption increase monotonically over time for as long as the agent does not involuntarily lose his job. If the agent is fired, his wealth accumulation is stopped and his consumption jumps downward. If the agent reaches the voluntary retirement threshold, his wealth and consumption reach their permanent, retirement levels smoothly. For any level of financial wealth the agent starts out with, we compute the planned duration of the agent's career, i.e., his time to planned retirement. Agents with lower initial wealth retire later.

We provide several comparative statics results. We show how the agent's optimal path of wealth accumulation and consumption prior to retirement depends on the risk of losing his job, on the value of leisure he obtains in retirement, and on the level of the rate of return paid by the asset in which the agent invests his savings. Higher job loss risk implies the agent saves more, consumes less, and retires faster. Lower utility of leisure implies the agent saves less, consumes more, and retires later. When the interest rate is higher, the agent retires with lower wealth and generally consumes more prior to retirement. In solving for the agent's optimal retirement rule, we discuss the option value of postponing retirement.

In addition, we discuss, in the context of our model, two standard properties of the solution to the optimal consumption and saving problem. We show that in the model with retirement, like in the standard model without retirement, the agent's marginal utility of consumption is a martingale, which means the conditional expected change in its value is always zero. We also review the result known as the permanent income hypothesis (PIH). Defined narrowly, PIH states that the agent chooses to consume exactly the income from his total wealth at all times. Total wealth consists of both financial and human wealth, where human wealth is defined as the expected present value of all labor income the agent is to earn in the future. With quadratic preferences, PIH holds in the standard model without retirement. We show that adding an endogenous retirement decision to the model does not overturn PIH.

We provide an elementary-level discussion of all dynamic optimization techniques involved in the analysis of our continuous-time model, thus making it accessible to a broad audience.

Related Literature

Our study is related to the literature on optimal consumption and saving decisions with fluctuating income and incomplete markets, and to the literature on the optimal timing of retirement.

The vast literature on the optimal saving problem with fluctuating income is summarized in Ljungqvist and Sargent (2004, Ch. 16). Classic studies of this problem, which include Friedman (1957), Bewley (1977), and Hall (1978), take the agent's stochastic income process as exogenous, which means they abstract from retirement. Chamberlain and Wilson (2000) allow for stochastic changes to the interest rate and show under weak conditions that optimal consumption diverges with probability one. Marcet, Obiols-Homs, and Weil (2007) extend the classic framework by including the agent's labor supply decision along the intensive margin. They show that with endogenous labor income the result of divergence of almost all consumption paths does not hold due to a wealth effect suppressing the agent's labor supply and thus eventually eliminating fluctuations in the agent's income. Our analysis is similar but allows for changes in labor supply along the extensive margin, i.e., it incorporates the retirement decision.

Similar to our analysis, Ljungqvist and Sargent (forthcoming) study an optimal consumption and saving problem with endogenous retirement. They focus on the impact of the curvature of the life cycle income profile on savings and the timing of retirement in a finite-horizon model in which all income shocks are unanticipated. Our model assumes a flat income profile in an infinite-horizon model in which the agent anticipates the risk in his income and responds to it.

Kingston (2000) and Farhi and Panageas (2007) study the optimal retirement timing decision combined with the problem of optimal saving and asset allocation prior to retirement, where available assets are one risky and one riskless asset, as in Merton (1971). They show that the option to delay retirement lets agents take on more risk than they would have chosen otherwise. In particular, Farhi and Panageas (2007) show that investors close to retirement may find it optimal to invest more heavily in stocks than those whose retirement is far off in the future. Our analysis is different as we do not consider a portfolio allocation problem in this article. Rather, we assume an incomplete market structure in which the riskless asset is the only vehicle for saving and wealth accumulation, as in the classic models of optimal consumption and saving decisions.

Our article is organized as follows. Section 1 presents our model. Section 2 discusses the optimal consumption pattern after retirement. Section 3 describes the optimal timing of retirement. Sections 4 and 5 study consumption and wealth accumulation prior to retirement.

Sections 6 and 7 provide comparative statics results with respect to several parameters of the model, with particular attention given to the job loss hazard parameter. Section 8 concludes. Appendix A contains proofs. Appendixes B and C discuss two extensions of the model.

1. MODEL

We will study the following partial equilibrium model in continuous time with a single agent. The agent consumes a single consumption good and leisure. The agent is initially employed. When employed, the agent earns a flow of labor income of $y > 0$ units of the consumption good per unit of time. The agent also consumes a flow of leisure of $l_W > 0$ units per unit of time. If the agent is not working, his labor income is zero but his flow of leisure is $l_R > l_W$. The agent's preferences over deterministic paths of consumption and leisure are represented by a standard utility function

$$\int_0^\infty e^{-rt} U(c_t, l_t) dt,$$

where c_t is consumption, $l_t \in \{l_W, l_R\}$ is leisure, and $r > 0$ is the agent's intertemporal rate of time preference.

While employed, the agent faces the risk of losing his job. If he loses his job, he never works again, which effectively means that losing one's job represents in our model involuntary retirement. The job loss shock arrives stochastically with a constant hazard rate $\lambda > 0$. That is, for any date t at which the agent is employed and for any $s > 0$, the probability that the agent will have not lost his job by date $t + s$ is $e^{-\lambda s}$.

In addition to losing his job involuntarily, the agent can quit. In this case, as well, the separation from employment is permanent, i.e., quitting means retiring. If he retires, the agent gives up the flow of labor income y and gains the flow of extra leisure $l_R - l_W > 0$.

At each point in time, the agent decides how much of his current income to consume and how much to save. There is only one asset in which the agent can invest his savings. It is a riskless asset with a constant rate of return equal to the agent's rate of time preference r. Denote the amount of the riskless asset held by the agent at date t, i.e., the agent's financial wealth at t, by W_t.

With these assumptions, the law of motion for the agent's financial wealth W_t is as follows. While working, the financial wealth changes according to

$$dW_t = (rW_t + y - c_t)dt. \tag{1}$$

Thus, for example, if the agent were to consume exactly his labor income while working, i.e., if $c_t = y$, then his financial wealth would grow exponentially at the rate of interest r. When not working (i.e., in retirement), the agent's wealth follows

$$dW_t = (rW_t - c_t)dt. \tag{2}$$

The agent maximizes

$$\mathbb{E}\left[\int_0^{\min\{\tau,\tau_f\}} e^{-rt}U(c_t, l_W)dt + \int_{\min\{\tau,\tau_f\}}^{\infty} e^{-rt}U(c_t, l_R)dt\right], \tag{3}$$

where τ is the agent's planned, voluntary retirement time, τ_f is the time he is forced into involuntary retirement, and the expectation $\mathbb{E}$ is taken over the realizations of the involuntary job loss shock. In particular, we will take the utility function to be separable in consumption and leisure:

$$\begin{aligned} U(c, l_W) &= u(c), \\ U(c, l_R) &= u(c) + \psi, \end{aligned}$$

where u is strictly increasing and a strictly concave utility of consumption and $\psi \geq 0$ is the utility of the extra leisure the agent enjoys in retirement. In this specification, the agent's lifetime utility (3) can be more simply written as

$$\mathbb{E}\left[\int_0^{\infty} e^{-rt}u(c_t)dt + e^{-r\min\{\tau,\tau_f\}}\frac{\psi}{r}\right].$$

2. OPTIMAL SAVING AND CONSUMPTION IN RETIREMENT

We start by discussing the agent's optimal use of savings in retirement. Because the return on the financial wealth held by the agent is equal to the agent's rate of time preference and the agent faces no uncertainty in retirement, it is natural to guess that in retirement the agent will keep assets constant, $dW_t = 0$, and consume his capital income, i.e., the return rW_t at all t. Thus, the natural guess is that if the agent retires with assets W_t, the maximum present value of total lifetime utility he can obtain after retirement, denoted by $V(W_t)$, is

$$V(W_t) = \frac{1}{r}u(rW_t) + \frac{\psi}{r}. \tag{4}$$

In the remainder of this section, we will use a standard dynamic programming argument to confirm that this guess is correct. In the process, we will derive an optimality condition on the value function

V—known as the Bellman equation—that will be useful when we discuss the agent's optimal consumption and saving behavior prior to retirement in the next section.

Following the dynamic programming approach, we take a small time interval $[t, t+h)$ and assume that from time $t+h$ onward the agent will apply the optimal saving and consumption policy, which is not known to us as of now. Given this assumption, we seek an optimal consumption rate c within the time interval $[t, t+h)$. Because h is small, we can consider c to be constant over the interval $[t, t+h)$. The true optimal consumption rate to be applied at time t, c_t, will be obtained by taking the limit as h goes to zero.

Because the agent follows an optimal consumption plan after $t+h$, the total discounted value he will obtain as of time $t+h$ will be $V(W_{t+h})$, where W_{t+h} is the amount of financial wealth the agent holds at $t+h$. For a given consumption rate c to be applied in $[t, t+h)$, the total discounted utility value the agent obtains as of time t is

$$\int_0^h e^{-rs}\left(u(c)+\psi\right)ds + e^{-rh}V(W_{t+h}). \tag{5}$$

Because this plan is a feasible consumption plan for an agent with nonnegative wealth, the maximal utility value $V(W_t)$ must be at least as large as the value of this plan, so for any c it is true that

$$V(W_t) \geq \int_0^h e^{-rs}\left(u(c)+\psi\right)ds + e^{-rh}V(W_{t+h}).$$

When h becomes arbitrarily small, the maximized value of the right-hand side of this expression approaches the value on the left-hand side, which we can write as

$$V(W_t) = \max_c \left\{\int_0^h e^{-rs}\left(u(c)+\psi\right)ds + e^{-rh}V(W_{t+h})\right\} \tag{6}$$

with h approaching zero. Since h is very small, we can replace the expression on the right-hand side of (6) with its first-order approximation. For a function f differentiable at some point t, for small h, we can approximate $f(t+h)$ with $f(t)+f'(t)h$. In this approximation, the first of the two terms in (5) equals

$$0 + \left(u(c)+\psi\right)h,$$

and the second term equals

$$V(W_t) + \left(-rV(W_t) + V'(W_t)\frac{dW_t}{dt}\right)h.$$

The value in (5) is therefore approximated by

$$V(W_t) + \big(u(c)+\psi - rV(W_t) + V'(W_t)(rW_t - c)\big)\,h,$$

where we have used the law of motion for assets in retirement (2). With this approximation, we can thus write (6) as

$$V(W_t) = \max_c \left\{ V(W_t) + \big(u(c) + \psi - rV(W_t) + V'(W_t)(rW_t - c)\big)\, h \right\}. \tag{7}$$

Dividing by h and simplifying terms, we obtain the following condition for the value function V:

$$rV(W_t) = \max_c \left\{ u(c) + \psi + V'(W_t)(rW_t - c) \right\}. \tag{8}$$

We will refer to this condition as the Bellman equation for the value function V. This equation shows how the agent's total utility $V(W_t)$ (converted to flow units by multiplying it by r) depends on current utility and the change in financial wealth. Higher c will increase the current utility flow $u(c) + \psi$ at the cost of lower saving $rW_t - c$. The marginal value of wealth $V'(W_t)$ shows how costly a change in saving is to the agent in utility terms. In choosing the consumption rate c the agent optimally balances this tradeoff between his utility from current consumption and his utility from future wealth.

Next, by differentiating the Bellman equation (8), we will obtain the optimal consumption policy function. Note that the Envelope Theorem lets us treat c as a constant in this differentiation. Indeed, differentiation gives us

$$rV'(W_t) = V''(W_t)(rW_t - c_t) + V'(W_t)r,$$

which simplifies to

$$0 = V''(W_t)(rW_t - c_t). \tag{9}$$

Assuming the second derivative V'' is nonzero, we divide both sides by $V''(W_t)$ to obtain

$$c_t = rW_t. \tag{10}$$

This confirms our guess that the optimal consumption policy for the agent in retirement is to consume the interest income from his financial assets at all t. Using this policy in the law of motion for wealth in retirement, (2), we confirm that $dW_t = 0$ and so assets and consumption remain constant in retirement. Substituting constant consumption $c_{t+s} = rW_t$ into the agent's utility function at all times $t + s$ following the retirement date t leads to the value function (4), confirming the guess we made at the beginning of this section.

We will also note that the first-order condition for the maximum on the right-hand side of (8) is

$$u'(c_t) = V'(W_t).$$

This condition, along with the policy function (10), lets us determine the marginal value of wealth in retirement as

$$V'(W_t) = u'(rW_t). \tag{11}$$

Clearly, the same result can be obtained by differentiating (4) directly.[1]

3. OPTIMAL RETIREMENT DECISION

In this section, we show that the optimal retirement policy for the agent is a threshold policy: The agent retires when his wealth reaches a specific threshold level. At this threshold, the marginal value of income the agent can earn if he works is exactly matched by the value of the extra leisure the agent can get if he retires.

In our analysis of the optimal voluntary retirement rule, we will use one intuitive property of the optimal pre-retirement wealth accumulation path. Namely, that the optimal wealth accumulation path is non-decreasing, i.e., the agent actually does save for retirement. That the agent will choose an increasing wealth accumulation path $\{W_t; t \geq 0\}$ prior to retirement is very intuitive in our model because the agent's labor income process is non-increasing and the return on savings is equal to the agent's rate of time preference. It is clear from (1) that W_t decreases only if $c_t > rW_t + y$, i.e., when the agent consumes more than his capital income rW_t and labor income y combined. Doing so clearly cannot be optimal for the agent given the labor income process the agent faces. The agent earns constant labor income $y > 0$ when he works and has no labor income after he quits or loses his job. In order to smooth consumption, the agent will want to save at least a part of his labor income for as long as he works, i.e., will choose $c_t \leq rW_t + y$ prior to retirement. In Section 5, we will characterize precisely what portion of y will be saved at each point in time. For now, we will just state that $c_t > rW_t + y$ is never optimal for the agent, and thus W_t is at least weakly increasing over time.

We now move on to the agent's optimal retirement decision. We will analyze this decision in two steps. First, we will compare the agent's value from retiring now, i.e., at some given time t, with the value from retiring a little later, i.e., at $t + h$, for a small $h > 0$. Then, we will argue that if the agent prefers to retire at t rather than retire at $t + h$ for a small h, then he also prefers to retire at t over retiring at any future date, which means the agent's overall optimal retirement time is t.

[1] Further, differentiating (4) twice, we have $V''(W_t) = ru''(rW_t) < 0$, which justifies the assumption of nonzero V'' we made when we divided (9) by $V''(W_t)$.

As before, we will use the first-order approximation for payoffs at $t+h$. In addition, we will discretize the involuntary job loss shock by assuming that if the agent loses his job by time $t+h$, this loss will occur only at $t+h$ and not earlier. With h approaching zero, these approximations will be sufficiently precise.

Suppose then that the agent is employed and has financial wealth W_t as of some time t. As we know from the previous section, the agents' value of retiring now is $V(W_t)$. The value of postponing retirement by a small amount of time h, denoted here by $V^h(W_t)$, is

$$V^h(W_t) = \max_c \left\{ \int_0^h e^{-rs} u(c) ds + e^{-rh} V(W_{t+h}) \right\}$$

with wealth following (1) between t and $t+h$, as the agent keeps working between t and $t+h$. Note that it does not matter if at $t+h$ the job loss shock happens or does not happen, because the agent is retiring at $t+h$ anyway. Since h is small, we use the first-order approximation and express $V^h(W_t)$ as

$$\begin{aligned} V^h(W_t) &= \max_c \left\{ V(W_t) + \left(u(c) - rV(W_t) + V'(W_t) \frac{dW_t}{dt} \right) h \right\} \\ &= \max_c \left\{ V(W_t) + \big(u(c) - rV(W_t) + V'(W_t)(rW_t + y - c) \big) h \right\}, \end{aligned}$$

where the second line uses (1).

Using the first-order approximation (7) for the value of retiring at t, $V(W_t)$, we have that postponing retirement by h is strictly preferred to retiring immediately, i.e., $V^h(W_t) > V(W_t)$, if and only if

$$\begin{aligned} & \max_c \left\{ V(W_t) + \big(u(c) - rV(W_t) + V'(W_t)(rW_t + y - c) \big) h \right\} \\ > \; & \max_c \left\{ V(W_t) + \big(u(c) + \psi - rV(W_t) + V'(W_t)(rW_t - c) \big) h \right\}. \end{aligned}$$

Dividing by h, simplifying, and taking terms that do not depend on c out of the maximization on each side, we have

$$\max_c \left\{ u(c) - V'(W_t)c \right\} + V'(W_t)y > \max_c \left\{ u(c) - V'(W_t)c \right\} + \psi.$$

Since the maximization problems on both sides of this inequality are the same, we simplify the above condition further to obtain

$$V'(W_t)y > \psi. \tag{12}$$

This says that whenever the utility flow from the additional leisure the agent can obtain by retiring is smaller than the utility he draws from the flow of his labor income, the agent will prefer to postpone retirement. From (11) we know that the agent's marginal value of wealth in retirement is $V'(W_t) = u'(rW_t)$. Thus, inequality (12) is equivalent to

$$\frac{\psi}{y} < u'(rW_t). \tag{13}$$

Let u'^{-1} denote the inverse function of u'. Since the right-hand side of (13) is strictly decreasing in W_t, it is true that this inequality holds for all $W_t < W^*$, where the threshold value W^* is given by

$$W^* = \frac{1}{r}u'^{-1}\left(\frac{\psi}{y}\right). \tag{14}$$

This means that postponing retirement (by at least a small instant) is preferred at all wealth levels W_t strictly smaller than W^*. The agent thus will not retire voluntarily with wealth $W_t < W^*$. Intuitively, for as long as his wealth is below W^*, by continuing to postpone retirement, the agent obtains a larger current flow return (his labor income is more valuable than the leisure forgone to obtain it) and retains the option to retire later.

Now that we know the agent will not retire with wealth smaller than W^*, we should ask if the agent will choose to retire as soon as his wealth reaches W^*. We know already that the agent with wealth W_t equal to or larger than W^* prefers to retire at t over retiring a bit later. But what about the possibility of retiring much later? Does $W_t \geq W^*$ also mean that the agent prefers to retire at date t rather than at any future date $T > t$? The answer is yes because, as we argued earlier in this section, the time path of wealth the agent chooses is never decreasing. Indeed, suppose the agent's wealth as of t satisfies $W_t \geq W^*$, but he does not retire until some later date $T > t$. Because the path of wealth is non-decreasing, $W_s \geq W^*$ at all dates s in $t \leq s \leq T$. In particular, for a small $h > 0$, at date $s = T - h$, the agent's wealth is greater than or equal to W^*, so, by our previous argument, the agent prefers to retire at $T - h$ rather than wait until T. Because his wealth is not smaller than W^* at $T - 2h$, as well, the agent will prefer to retire at $T - 2h$ rather than at $T - h$. Extending this reasoning backward in time all the way back to date t shows that the agent's overall preferred retirement rule is to retire as soon as his wealth reaches W^*.

In sum, the optimal retirement rule takes on a threshold form. The agent chooses to postpone retirement for as long as his wealth is below the threshold W^* and retire immediately when his wealth reaches W^*. It is worth noting in (14) that the optimal wealth threshold W^* increases in labor income y, decreases in the value of leisure ψ, and does not depend on the intensity of the job loss risk λ. If $\psi = 0$, i.e., if working is not costly to the agent in terms of forgone leisure at all, then $W^* = \infty$. In this case, the agent never chooses to retire voluntarily.

The Option Value of Postponing Retirement

Because retirement is permanent in our model, when the agent retires he loses the option of working at later dates. The threshold retirement rule we derived tells us, however, that the value of this option is zero for the agent in the problem we study.

In general, a one-time, irreversible action has a positive option value for an agent if he is willing to forgo an immediate benefit that the action can produce in order to retain the option of taking the action in the future.[2] In our model, the agent retires as soon as the current flow return from doing so turns positive, i.e., when the value of the flow of leisure, ψ, becomes as large as the value of the flow of labor income y, $V'(W_t)y$. The agent is not willing to delay retirement beyond that point because once wealth reaches W^* the agent will continue to prefer the flow of leisure ψ over the flow of his labor income y at all future times in all possible realizations of uncertainty he faces. In fact, once the agent retires with wealth $W_t \geq W^*$, there is no realization of uncertainty in which he might want to go back to working, even if he could return.

The value of having the option to work in the future that the agent gives up by retiring would in our model be positive if the parameters determining the threshold wealth level W^* could change in a way that increases the value of working relative to the value of consuming leisure. In particular, the value of this option would be positive if the agent could receive a positive income shock increasing the level of his labor income y, or a taste shock decreasing the utility of leisure ψ, or a taste shock increasing the agent's marginal utility of consumption u', or a shock destroying a part of the agent's financial wealth W_t. In Appendix C, we discuss this point in more detail, focusing on the possibility of an increase in labor income y.

4. CONSUMPTION, SAVING, AND WEALTH ACCUMULATION PRIOR TO RETIREMENT

In this section, we study the agent's optimal saving and consumption decisions prior to retirement, i.e., when his wealth is strictly less than W^*. The guess-and-verify method we used earlier to solve for optimal consumption in retirement will not work here because wealth and consumption have nontrivial dynamics prior to retirement. In order to

[2] For example, it is often optimal for a business owner to keep her business open for some time after it begins to make a loss (a flow of negative profit). The option for the profits that the business may generate in the future has a value that keeps the owner from shutting the business down as soon as the current profit flow turns negative (see Leland [1994]). Pindyck (1991) discusses the option value of undertaking a one-time investment in a stochastic environment.

study these dynamics, we will derive intertemporal optimality conditions leading to a dynamic system in wealth and consumption. We will then use standard methods to analyze this system.

Bellman Equation

Let us denote by $J(W_t)$ the maximal discounted expected utility value a working agent can obtain given his wealth W_t. Since retiring immediately is optimal when $W_t \geq W^*$, we have $J(W_t) = V(W_t)$ for all $W_t \geq W^*$. Since not retiring is strictly preferred by the agent when $W_t < W^*$, we have $J(W_t) > V(W_t)$ for all $W_t < W^*$. We look now to learn more about $J(W_t)$ for $W_t < W^*$. We proceed by deriving the Bellman equation for J analogous to the Bellman equation for V we derived earlier.

Take a small $h > 0$ and assume that an agent who works at t and holds financial wealth W_t chooses to consume at some constant rate c inside the time interval $[t, t+h)$. In addition, assume that if the agent wants to quit inside $(t, t+h)$, he will do so only at $t+h$. Likewise, assume that if the agent loses his job involuntarily during this short period of time, this will happen only at the end of the period, i.e., at date $t+h$. As before, these assumptions will be innocuous when we take the limit with h going to zero. Following the dynamic programming approach, we suppose that from time $t+h$ onward the agent applies an optimal (to us yet unknown) consumption and saving policy. The total utility value the agent obtains by following this strategy with some fixed consumption rate c is

$$\int_t^{t+h} e^{-rs} u(c) ds + e^{-rh} \left[e^{-\lambda h} J(W_{t+h}) + (1 - e^{-\lambda h}) V(W_{t+h}) \right]. \quad (15)$$

The term in square brackets represents the expectation of the value the agent will draw at time $t+h$. With probability $e^{-\lambda h}$ he does not lose his job as of $t+h$ and $J(W_{t+h})$ represents the continuation value he obtains at that time. With probability $1 - e^{-\lambda h}$ he loses his job, and thus the continuation value he obtains at $t+h$ is the retirement value $V(W_{t+h})$.

With the optimal choice of c, the value in (15) approaches the overall maximal value the agent can obtain, $J(W_t)$, which we write as

$$\begin{aligned} J(W_t) \quad = \quad & \max_c \Big\{ \textstyle\int_t^{t+h} e^{-rs} u(c) ds \\ & + e^{-rh} \Big[e^{-\lambda h} J(W_{t+h}) + (1 - e^{-\lambda h}) V(W_{t+h}) \Big] \Big\} \quad (16) \end{aligned}$$

with h approaching zero. Since h is very small, we can apply the first-order approximation to the value in (15) and write it as

$$J(W_t) + \big(u(c) - (r+\lambda)J(W_t) + J'(W_t)(rW_t + y - c) + \lambda V(W_t)\big)h.$$

Using this approximation in (16), we have

$$\begin{aligned} J(W_t) &= \max_c \big\{J(W_t) + \big(u(c) - (r+\lambda)J(W_t) + \\ &\quad + J'(W_t)(rW_t + y - c) + \lambda V(W_t)\big)h\big\}. \end{aligned}$$

Dividing by h and simplifying terms, we get the Bellman equation for J:

$$(r+\lambda)J(W_t) = \max_c \big\{u(c) + J'(W_t)(rW_t + y - c) + \lambda V(W_t)\big\}. \quad (17)$$

To compare it with the Bellman equation for V, (8), let us rewrite (17) as

$$rJ(W_t) = \max_c \big\{u(c) + J'(W_t)(rW_t + y - c)\big\} - \lambda\left(J(W_t) - V(W_t)\right). \quad (18)$$

Bellman equations (8) and (18) differ in three ways. First, the trade-off between consumption and saving is different, as prior to retirement the agent earns the stream of income y. Second, the level of J is also influenced by the lower flow of leisure prior to retirement. These two differences are reflected in the expression inside the maximization with respect to c in (18). Third, (18) contains an extra term, $-\lambda\left(J(W_t) - V(W_t)\right)$, that reflects the possibility of the agent's involuntarily losing his job. In this term, λ is the intensity with which the agent loses his job and $J(W_t) - V(W_t)$ is the loss of value that occurs in that event.

Euler Equation

As before, we use the envelope and first-order conditions associated with the Bellman equation. Using the Envelope Theorem in differentiation of the Bellman equation (17) yields

$$(r+\lambda)J'(W_t) = J''(W_t)(rW_t + y - c_t) + J'(W_t)r + \lambda V'(W_t).$$

Simplifying terms and rearranging, we get

$$\lambda\left(J'(W_t) - V'(W_t)\right) = J''(W_t)\left(rW_t + y - c_t\right). \quad (19)$$

Unlike in the post-retirement problem we studied earlier, in the pre-retirement problem the envelope condition (19) does not by itself determine the optimal consumption rule. However, it gives us an important intertemporal optimality condition for consumption known as the Euler equation. To derive it, we use the chain rule to express the time derivative of $J'(W_t)$ as

$$\begin{aligned}\frac{dJ'(W_t)}{dt} &= J''(W_t)\frac{dW_t}{dt} \\ &= J''(W_t)(rW_t + y - c_t),\end{aligned}$$

where the second equality uses (1). This lets us write (19) as

$$\frac{dJ'(W_t)}{dt} = \lambda\left(J'(W_t) - V'(W_t)\right).$$

Next, we use the first-order condition in the maximization problem in the Bellman equation (17),

$$u'(c_t) = J'(W_t), \tag{20}$$

to write the above as

$$\frac{du'(c_t)}{dt} = \lambda\left(u'(c_t) - V'(W_t)\right).$$

Finally, we use (11) to eliminate V' from the above equation and express it purely in terms of the marginal utility of consumption:

$$\frac{du'(c_t)}{dt} = \lambda\left(u'(c_t) - u'(rW_t)\right). \tag{21}$$

This is the Euler equation for consumption prior to retirement. It shows how the marginal utility of consumption changes along an optimal path of consumption and financial wealth accumulation prior to retirement.[3]

Martingale Property

Before we use the Euler equation to study optimal consumption and asset accumulation, let us discuss an implication of the Euler equation known as the martingale property of marginal utility. As studied by Hall (1978) and many others, (21) implies that at all times prior to retirement the expected change in marginal utility of consumption is zero, i.e., marginal utility of consumption is a so-called martingale.[4] In discrete-time models that are most commonly used in the literature, the Euler equation takes the familiar form of $u'(c_t) = \mathbb{E}_t\left[u'(c_{t+1})\right]$ at all t, where $\mathbb{E}_t\left[\cdot\right]$ is the conditional expectation operator. In discrete time, it is thus easy to see that the expected change in u' is zero. In continuous time, the martingale property is slightly less self-evident but can still be seen as follows.

[3] Note that, trivially, the Euler equation also holds after retirement.

[4] Because consumption is constant in retirement, marginal utility of consumption is trivially also a martingale after the agent retires, voluntarily or not.

Take a small $h > 0$ and a date t at which the agent is not retired. In the time interval $[t, t+h)$, the agent will be hit with the job loss shock with probability $1-e^{-\lambda h}$, which will cause his marginal utility at $t+h$ to change (jump) by $u'(rW_{t+h})-u'(c_{t+h})$. With probability $e^{-\lambda h}$, the agent will not lose his job, in which case the change in his marginal utility over the time interval $[t, t+h)$ will be simply $u'(c_{t+h}) - u'(c_t)$. Marginal utility is a martingale when the average (i.e., expected) value of these two changes is zero, i.e., when

$$(1 - e^{-\lambda h}) \left(u'(rW_{t+h}) - u'(c_{t+h})\right) + e^{-\lambda h} \left(u'(c_{t+h}) - u'(c_t)\right) = 0.$$

Rearranging this condition, we have

$$\begin{aligned} u'(c_{t+h}) - u'(c_t) &= (e^{\lambda h} - 1) \left(u'(c_{t+h}) - u'(rW_{t+h})\right) \\ &= \lambda h \left(u'(c_{t+h}) - u'(rW_{t+h})\right), \end{aligned}$$

where the second equality uses the linear approximation $e^{\lambda h} = 1 + \lambda h$. Dividing by h and taking formally the limit as $h \to 0$, we get the Euler equation (21). Thus, the Euler equation (21) says exactly that the time trend $du'(c_t)/dt$ in marginal utility along the path that consumption follows conditional on the job loss shock not occurring is the negative of the jump in marginal utility that occurs if the agent loses his job, $u'(rW_t)-u'(c_t)$, times the intensity of the job loss λ. This trend exactly offsets the jump-induced change in marginal utility, making the overall expected change in marginal utility zero, i.e., marginal utility indeed is a martingale.[5]

Dynamic Analysis

As we saw earlier, consumption and financial wealth have trivial dynamics in retirement: Both remain constant over time. Prior to retirement, however, wealth and consumption do change over time. We will now use the Euler equation (21) and the law of motion for wealth (1) to study the dynamics of wealth and consumption prior to retirement. To do this, we use the chain rule

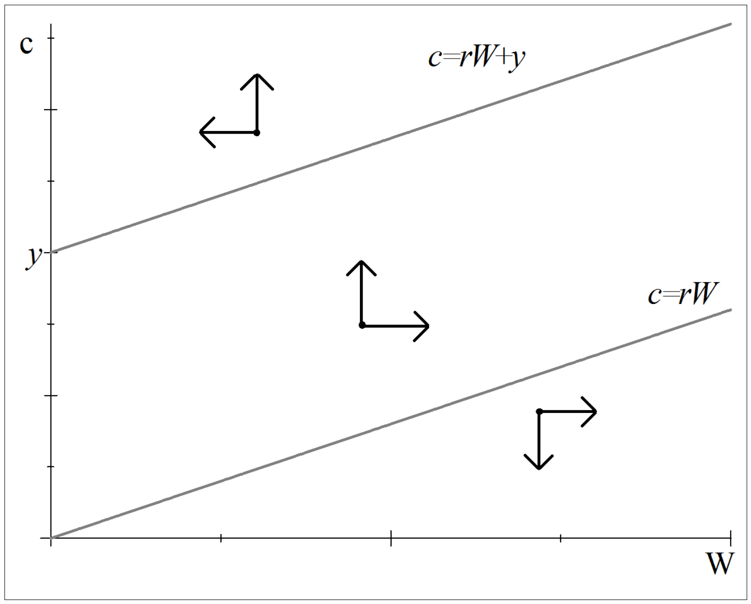
$$\frac{du'(c_t)}{dt} = u''(c_t)\frac{dc_t}{dt}$$

and the strict concavity of u, implying $u'' \neq 0$, to express the Euler equation (21) as

$$\frac{dc_t}{dt} = \lambda \frac{u'(c_t) - u'(rW_t)}{u''(c_t)}. \tag{22}$$

[5] In this respect, the marginal utility process is in our model similar to a compensated Poisson process. See Problem 1.3.4 in Karatzas and Shreve (1997).

Figure 1 Phase Diagram for (W,c)



Together with the law of motion for financial wealth W_t, (1), this gives us a dynamic system describing the evolution of consumption and wealth prior to retirement. In the rest of this subsection, we will study qualitative properties of this system. We will use a phase diagram to describe the shape of the time paths in the plane (W, c) that satisfy the differential equations (1) and (22). Any such path is called a solution to the system (1), (22), and there are an infinite number of them (every point in the domain for (W, c) belongs to one solution). The solutions represent all paths of consumption and wealth accumulation the agent might want to follow while working that are consistent with intertemporal optimization. That is, any path that is not a solution to (1), (22) is not optimal for the agent. In order to select the optimal path from among all solutions to this system, a boundary condition is needed. In standard infinite-horizon analysis, the transversality condition serves this role. In our model with endogenous retirement, this condition will be provided by the optimal voluntary retirement rule we obtained in the previous section.

The phase diagram for the system of differential equations (1), (22) is shown in Figure 1. It provides a graphical representation of the directions in which the system (W_t, c_t) moves along all possible solution paths. In Figure 1, these directions are marked by horizontal and vertical arrows. These arrows are determined as follows.

From (22) we see that along a solution path consumption will increase over time, i.e., $dc_t/dt > 0$, if and only if $u'(c_t) - u'(rW_t) < 0$, i.e., if and only if $c_t > rW_t$. The line $c = rW$ therefore divides the state space (W, c) into two regions: one in which consumption grows over time (the region above this line), and one in which it decreases over time (the region below it). Similarly, we have from (1) that W_t grows over time if and only if $c_t < rW_t + y$. Therefore, the line $c = rW + y$ divides the state space (W, c) into a region of wealth growth (below this line) and a region of wealth decline (above it). Since the two lines are parallel, we see that there are three regions in the state space (W, c) differentiated by distinct dynamic properties of the system (W_t, c_t). Above the line $c = rW + y$, wealth declines and consumption grows. In the band $rW < c < rW + y$, both wealth and consumption grow. Below the line $c = rW$, wealth grows and consumption declines.

The qualitative conclusions we can obtain from the phase diagram are as follows. Inside the band $rW < c < rW + y$, solution paths increase in both the c and the W direction and fall into one of the following three types. Paths of the first type will reach the upper straight line $c = rW + y$, where they bend backward as W_t begins to decrease while c_t continues to increase. Paths of the second type will reach the lower straight line $c = rW$, where they bend downward with c_t declining and W_t continuing to increase. Note that none of the paths of the first or second type return to the band $rW < c < rW + y$ once they leave it. Paths of the third type will stay inside the band $rW < c < rW + y$ forever.

Further characterization of the solution paths can be obtained analytically in the special case in which the Euler equation (22) is linear or numerically in other cases. In the remainder of this article, we will focus on the case with a linear Euler equation and discuss analytical solutions. As we will see in the next section, the Euler equation is linear when the utility function u is quadratic. In Appendix B, we briefly discuss how the results change for other utility functions, in particular for preferences exhibiting constant relative risk aversion (CRRA).

5. EXACT SOLUTION WITH LINEAR EULER EQUATION

We specialize the utility function to

$$u(c) = -\frac{1}{2}(c - B)^2$$

and restrict its domain to $c \leq B$. Under this specification, marginal utility is linear and therefore the Euler equation (22) is linear in consumption as well as in wealth:

$$\frac{dc_t}{dt} = \lambda \left(c_t - rW_t \right). \tag{23}$$

The system of differential equations (1), (23) is now linear and can be solved in closed form. In particular, we have the following lemma providing analytical expressions for all solution paths of the system (1), (23).

Lemma 1 *Let $\{(W_t, c_t); t \geq 0\}$ be a solution path. If there exists τ such that $c_\tau = rW_\tau + y$, then*

$$c_t = rW_t + \frac{ry}{r+\lambda}\left(1 - e^{-(r+\lambda)(\tau - t)}\right) + ye^{-(r+\lambda)(\tau-t)}. \tag{24}$$

If there exists τ such that $c_\tau = rW_\tau$, then

$$c_t = rW_t + \frac{ry}{r+\lambda}\left(1 - e^{-(r+\lambda)(\tau - t)}\right). \tag{25}$$

Otherwise,

$$c_t = rW_t + \frac{ry}{r+\lambda}. \tag{26}$$

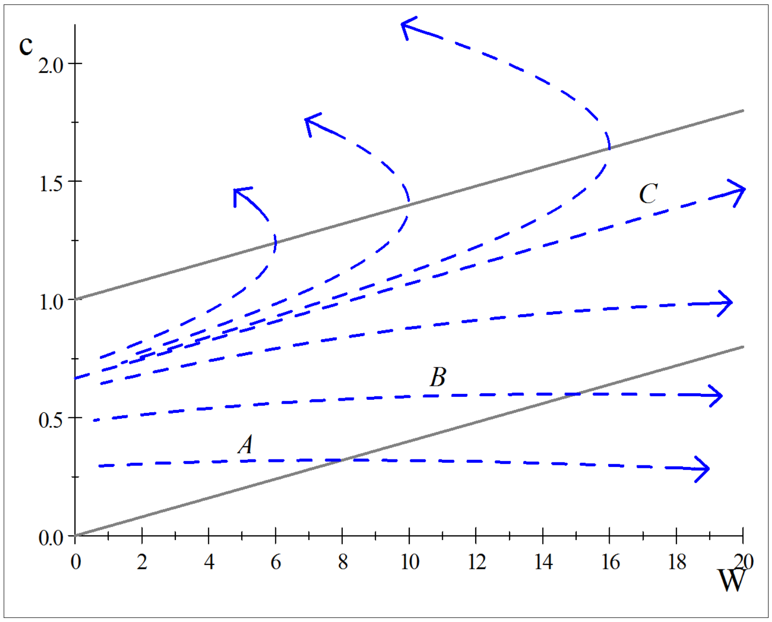
Proof. In Appendix A. ■

Figure 2 plots several sample solution paths $\{(W_t, c_t); t \geq 0\}$ of the three types given in the above lemma. Solution paths (24) bend backward with wealth declining over time at all dates $t > \tau$, where τ is such that $c_\tau = rW_\tau + y$. None of these solution paths will be optimal for the agent because, as we saw earlier, it is never optimal for the agent to see his financial wealth decrease while he is saving for retirement. Along all solution paths (25) and (26) wealth is increasing. These paths, therefore, are our candidates for the optimal path of consumption and saving prior to retirement.

The Optimal Accumulation Path

As we saw in Section 3, the agent's retirement decision is determined by a simple wealth threshold rule. The agent retires as soon as his wealth reaches the level W^*. The threshold W^* depends on the parameters r, ψ, and y, as shown in (14). At retirement (and forever after), the agent's optimal level of consumption is $c_t = rW^*$. The Euler equation (23) and the wealth accumulation equation (1) tell us that prior to retirement the agent follows one of the non-backward-bending paths depicted in Figure 2. For a given value of W^*, which one of these paths will the agent follow?

Figure 2 Solution Paths with Quadratic Utility



Notes: Parameters used in this plot: $y = 1$, $r = 0.04$, $\lambda = 0.02$.

Since the agent's wealth and consumption remain constant in retirement, in Figure 2 the evolution of wealth and consumption after voluntary retirement is represented by a single point for each threshold value W^*. That point is (W^*, rW^*). Thus, once the agent retires, the time path of his wealth and consumption is absorbed at (W^*, rW^*). It is easy to see in Figure 2 that for each value $W^* \geq 0$ there is a unique solution path $\{(W_t, c_t); t \geq 0\}$ leading to the point (W^*, rW^*). That path is the optimal path for the agent whose retirement wealth threshold is W^*. Why this path? Because all other paths would imply a jump in consumption at retirement, which the agent wants to avoid. Because his utility function is concave, the agent prefers a smooth consumption path with no jump at retirement. The level of consumption in voluntary retirement, rW^*, thus determines the optimal accumulation path the agent follows prior to retirement. It is the single path that intersects the line $c = rW$ at $W = W^*$.

For example, the solution path labelled A in Figure 2 crosses the line $c = rW$ at $W = 8$. Thus, this solution path is optimal for the agent whose desired retirement wealth is $W^* = 8$. Likewise, the solution path B is optimal for the agent whose desired retirement wealth is $W^* = 15$. The solution path C follows a straight line parallel to the line $c = rW$ and therefore never crosses it. This solution path is optimal for the agent whose retirement threshold is $W^* = \infty$, which means the agent plans to never retire voluntarily. From the formula (14) we see that $W^* = \infty$ when $\psi = 0$, i.e., when the agent does not at all value the extra leisure he can obtain by retiring. Since the value of the extra leisure is zero for this agent, it is natural that he never chooses to retire. This is the case studied in standard infinite-horizon models of optimal saving and consumption decisions, e.g., Ljungqvist and Sargent (2004, Ch. 16).

Note that the argument implying that the agent's preferred path of wealth and consumption is the one that leads to the point (W^*, rW^*) does not use the assumption of quadratic preferences. Rather, this argument is based on the agent's preference for smooth consumption, and so it applies to any concave utility function u. Thus, although the shape of the optimal path of wealth accumulation and consumption $\{(W_t, c_t); t \geq 0\}$ will in general not be the same as that presented in Figure 2, it will be true for any concave utility function that the optimal accumulation path is the unique solution path that leads to the point (W^*, rW^*).[6]

Also, the phase diagram in Figure 1, which works for any concave u, shows that the only way for a solution path to approach (W^*, rW^*) is through the middle band $rW < c < rW + y$ of the state space (W, c), where W_t and c_t are both strictly increasing over time. This confirms the validity of the assumption we made in Section 2 about wealth following an increasing path prior to retirement.

Planned Retirement and Optimal Saving Rate

Figure 2 provides a clear illustration of the following point. When the option to retire is added to the standard, infinite-lived-agent model of optimal consumption and saving, the model's prediction on the optimal amount of saving unambiguously increases.

In its textbook version (see Ljungqvist and Sargent [2004]), the standard model of optimal consumption and saving decisions abstracts

[6] See Appendix B for the case of the CRRA utility function.

from retirement. Labor income fluctuates stochastically, but the agent does not have an option to retire and end his flow of labor income altogether. The model we consider in this article assumes a particularly simple form of stochastic income fluctuations (labor income is a step function initially positive and jumping down to zero at a random date τ_f), but allows for endogenous retirement.

In the special case with $\psi = 0$, our model is a version of the standard model with no retirement. As we saw earlier, when $\psi = 0$ the agent never retires voluntarily, so his labor income effectively follows an exogenous process, as in the standard model. From (26) we have in that case that the optimal fraction of labor income y to be saved by the agent is

$$\begin{aligned} \frac{dW_t/dt}{y} &= \frac{rW_t + y - c_t}{y} \\ &= 1 - \frac{r}{r+\lambda} \\ &= \frac{\lambda}{r+\lambda}. \end{aligned}$$

Thus, the standard model without retirement would predict $\frac{\lambda}{r+\lambda}$ as the agent's optimal rate of saving out of labor income. With positive utility of leisure, $\psi > 0$, our model predicts voluntary retirement in finite time τ as well as a higher optimal rate of saving prior to retirement. From (25) we have

$$\begin{aligned} \frac{dW_t/dt}{y} &= \frac{rW_t + y - c_t}{y} \\ &= 1 - \frac{r}{r+\lambda}\left(1 - e^{-(r+\lambda)(\tau - t)}\right) \\ &> \frac{\lambda}{r+\lambda}, \end{aligned}$$

where the strict inequality follows from the agent's time to retirement being finite, i.e., $\tau - t < \infty$. Given that people do save for retirement in reality, models that disregard retirement underpredict the optimal rate of saving. Figure 2 shows this very clearly: The infinite-horizon solution path that runs parallel to the line $c = rW$ is everywhere above all solution paths that cross this line.

That the optimal saving rate should be higher when agents save for retirement is of course very intuitive. With retirement, the agent's labor income is more front-loaded relative to the case without retirement. To smooth this front-loading out, the agent saves more. Our analysis lets us see this point clearly in Figure 2.[7]

Time to Retirement

As we see in (25), the agent's optimal consumption at time t depends on the agent's current wealth W_t and the amount of time left before his planned retirement, $\tau - t$. The agent's target retirement wealth level W^* is given in (14). But how do we find the agent's target retirement time τ?

From the law of motion for wealth prior to retirement, (1), we have

$$W_\tau = W_t + \int_0^{\tau-t} dW_{t+s} = W_t + \int_0^{\tau-t} (y - (c_{t+s} - rW_{t+s}))ds.$$

Using the retirement condition $W_\tau = W^*$ and the consumption rule (25) we have

$$\begin{aligned} W^* &= W_t + \int_0^{\tau-t} \left(1 - \frac{r}{r+\lambda}(1 - e^{-(r+\lambda)s})\right) yds \\ &= W_t + \frac{\lambda}{r+\lambda}(\tau - t)\, y + \frac{r}{(r+\lambda)^2}\left(1 - e^{-(r+\lambda)(\tau-t)}\right)y. \quad (27) \end{aligned}$$

For any given values for r, λ, y, W^*, and W_t, this condition can be solved for the agent's planned time to retirement $\tau - t$. Because the right-hand side of (27) is increasing in both $\tau - t$ and W_t, the time to retirement is decreasing in current wealth.[8]

In sum, the dynamics of consumption and wealth accumulation are as follows. The agent determines his target retirement wealth level W^*, as in (14). Then the agent follows the unique wealth accumulation and consumption path $\{(W_t, c_t); t \geq 0\}$ in Figure 2 that leads to the point (W^*, rW^*). How far away from the retirement point the agent starts on this path depends on his initial wealth W_0. Unless he loses his job before reaching wealth W^*, the agent retires voluntarily as soon as his wealth attains W^*. After retirement, he consumes at the constant rate rW^* and his financial wealth remains constant at W^*. Thus, the solution path the agent follows in Figure 2 is absorbed at the point (W^*, rW^*). If the agent is forced into involuntary retirement at some date $\tau_f < \tau$, i.e., when his wealth is $W_{\tau_f} < W^*$, his consumption jumps down at τ_f from his preferred accumulation path to the point $(W_{\tau_f}, rW_{\tau_f})$, and is absorbed there. That is, consumption stays constant in retirement at

[7] The same is true in the case of CRRA preferences we discuss in Appendix B. See Figure 7.

[8] This also confirms that wealth grows over time while the agent is working.

the level $rW_{\tau_f} < rW^*$ and financial wealth stays constant at $W_{\tau_f} < W^*$.

Permanent Income Hypothesis

It is well known, see Ljungqvist and Sargent (2004, Ch. 16), that with quadratic preferences the optimal saving and consumption rule satisfies the permanent income hypothesis (PIH). Under PIH, it is optimal for the agent at each point in time to consume simply the income from his total wealth, where total wealth includes financial wealth and human capital. Human capital of an agent is defined as the present value of all the labor income that the agent is yet to earn. Thus, permanent income has two components: the income from currently held financial wealth and the income from currently held human capital.

That PIH holds in our model is most clearly evident in (26), i.e., in the case with $\psi = 0$ in which the agent never retires voluntarily. If the agent's stock of financial wealth is W_t, his permanent income from it is rW_t because, as we saw earlier, if the agent consumes rW_t, he never depletes his financial wealth and therefore is able to maintain this consumption forever. If the agent is working at t, the expected present value of his future labor income is

$$\mathbb{E}\left[\int_0^{\tau_f} e^{-rs} y ds\right] = \frac{y}{r+\lambda}.$$

Permanent income from human capital $y/(r+\lambda)$ is $ry/(r+\lambda)$ because this is the perpetual flow equivalent of stock $y/(r+\lambda)$. According to PIH, with financial wealth W_t and with human capital $y/(r+\lambda)$, the agent's consumption at t should be $r(W_t + \frac{y}{r+\lambda})$, which it is, as we see in (26).

The agent's optimal rule for consumption and saving obeys PIH also when he chooses to voluntarily retire at a future date τ. In this case, the agent's human capital as of $t < \tau$ is

$$\mathbb{E}\left[\int_0^{\min\{\tau_f,\tau\}} e^{-rs} y ds\right] = \frac{y}{r+\lambda}\left(1 - e^{-(r+\lambda)(\tau-t)}\right). \qquad (28)$$

Thus, (25) is consistent with PIH because the agent in this case as well consumes exactly the return on his financial and human capital at all times. Note that the value in (28) is less than $y/(r+\lambda)$ because a part of expected future income is lost due to the agent's planned retirement at τ. The closer t is to τ, the lower the agent's human capital. Because the agent saves at all $t < \tau$, however, his financial wealth W_t grows as t gets closer to τ. It fact, financial wealth grows faster than human capital declines, and so the agent's permanent consumption increases over time for as long as the agent does not lose his job. As wealth approaches W^*, the agent's human capital goes down to zero smoothly, and his consumption increases smoothly to rW^*. If the agent loses his job involuntarily at some date τ_f before his financial wealth reaches W^*, the agent's human capital discontinuously jumps down to zero and his permanent consumption jumps down to just the return on his financial assets, rW_{τ_f}.

6. RETIREMENT SAVING AND THE JOB LOSS RISK

In this section, we study the dependence of the optimal consumption and saving plan on the job loss rate λ.

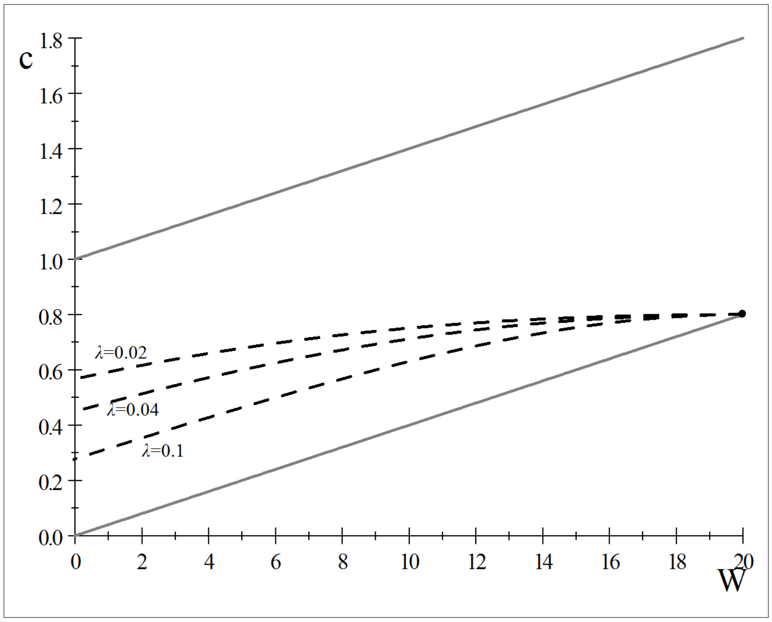
Proposition 1 *At any $W_t < W^*$, the larger the job loss intensity λ, the lower consumption c_t, the higher the wealth accumulation rate dW_t/dt, and the shorter the time to planned retirement $\tau - t$. If $\lambda \to \infty$, then $c_t \to rW_t$, $dW_t/dt \to y$, and $\tau - t \to (W^* - W_t)/y$.*

Proof. In Appendix A. ■

This proposition shows that if we compare two agents identical in all respects (same wealth, same income) except for the job loss rate λ, the agent with larger job loss risk will consume less and save more than the other agent. Intuitively, the agent with higher λ holds less human capital than the agent whose λ is lower. The labor income flow rate y, the same for both agents, therefore, is higher relative to total wealth for the agent with higher λ, and so he will save a larger portion of y than the other agent. In other words, labor income y is less permanent for the agent with higher λ, so intertemporal consumption smoothing implies he will save more. Figure 3 illustrates this point by plotting optimal paths for consumption and wealth for several values of λ.

This comparative statics result can be interpreted as showing the agent's response to a completely unanticipated shock to the job loss risk the agent faces in our model. Under the parametrization used in Figure 3, if $\lambda = 0.02$, the agent will follow the highest of the three accumulation paths plotted in that figure. If at some point prior to retirement the intensity parameter λ jumps to 0.1, the agent will switch at that point to the lowest of the three paths. This means that his saving rate will increase and consumption will decrease without any change to his current income. This example illustrates a response of

Figure 3 Optimal Accumulation Paths for Three Values of the Job Loss Rate



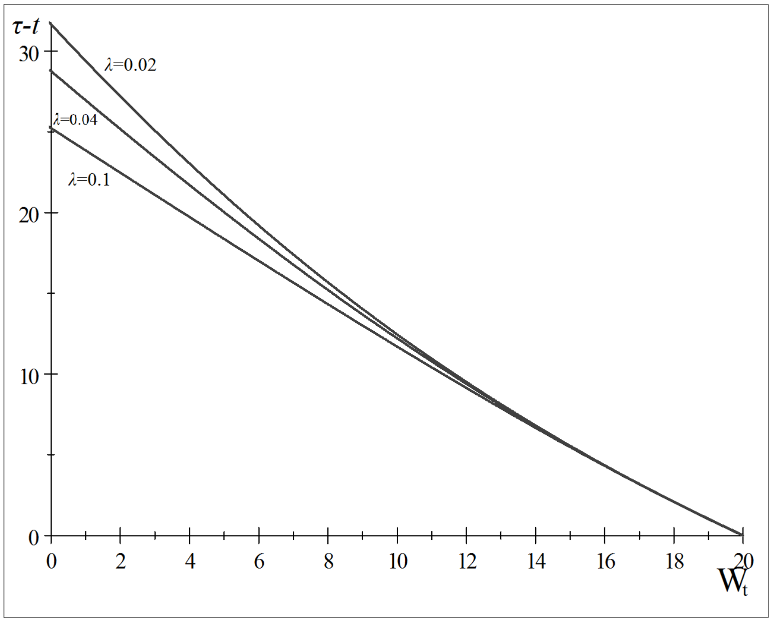
Notes: Other parameters used in this plot: $r = 0.04$, $y = 1$, ψ such that $W^* = 20$.

optimal consumption to a change in the expectations the agent holds about the future.[9]

If λ is very large, then, as Proposition 1 shows, the agent saves close to 100 percent of his labor income y and consumes close to rW_t. This again is intuitive, as when λ is large, the agent's human capital is close to zero and financial wealth constitutes the bulk of his total wealth. The level of permanent consumption he can afford is thus close to the level he could maintain if he had lost his job already, which with assets W_t is exactly rW_t.

[9] The discussion in this paragraph assumes that the agent does not anticipate that λ could jump, and that once it does jump, the agent firmly expects it to never jump again. Clearly, this is an oversimplification. We can expect, however, that our conclusion here continues to hold when the jumps in λ are anticipated. That is, although the shape of the accumulation paths in Figure 3 must be adjusted, we expect consumption to decline when λ increases in a model in which changes in λ are anticipated by the agent.

Figure 4 Dependence of Time to Retirement on the Job Loss Rate



Notes: Parameters used in this plot are the same as those in Figure 3.

Proposition 1 also shows that conditional on not losing the job before the planned retirement date τ, the agent with larger λ will reach his desired retirement wealth level W^* faster. Note that the theoretical limit with $\lambda \to \infty$ of the time to retirement, $(W^* - W_t)/y$, is consistent with the agent saving 100 percent of his labor income and living only off his asset income already before retirement.

Figure 4 plots the planned time to retirement $\tau - t$ against wealth W_t for several values of λ. In the example presented in that figure, we have $r = 0.04$, which makes one unit of time correspond roughly to one year. Annual labor income y is normalized to 1, and $W^* = 20$, which means that the agent wants to retire as soon as his stock of wealth reaches the equivalent of 20 years of labor income. With $\lambda = 0.02$, meaning the event of involuntary and permanent job loss on average occurs once in 50 years, the agent who starts out with zero initial wealth plans to retire after about 32 years. With $\lambda = 0.04$, i.e., when involuntary retirement is a once-in-a-quarter-century event, the agent plans to retire after roughly 28.5 years. With $\lambda = 0.01$, the permanent job loss shock becomes a once-in-a-decade event in expectation. In this case, the agent plans to retire after 25 years. These numbers illustrate the fact that the agent can only partially insure himself against the permanent job loss shock in our model. With $\lambda = 0.02$, the probability that an agent with zero wealth reaches voluntary retirement is $e^{-0.02*32}$, which equals roughly 53 percent. For an agent with the same initial wealth but with $\lambda = 0.1$, this chance is only $e^{-0.1*25}$, i.e., about 8 percent.

Differentiating with respect to λ the expressions for optimal consumption in (25) and (26), it is easy to check that the response of c_t to a given change in λ is stronger the longer the agent's planned time to retirement $\tau - t$. In particular, the response of consumption to changes in the job loss risk is the strongest in the case of $\psi = 0$, where the agent plans to never retire voluntarily. This result is very intuitive given that fast planned retirement means human capital is a small portion of the agent's total wealth.

7. ADDITIONAL COMPARATIVE STATICS RESULTS

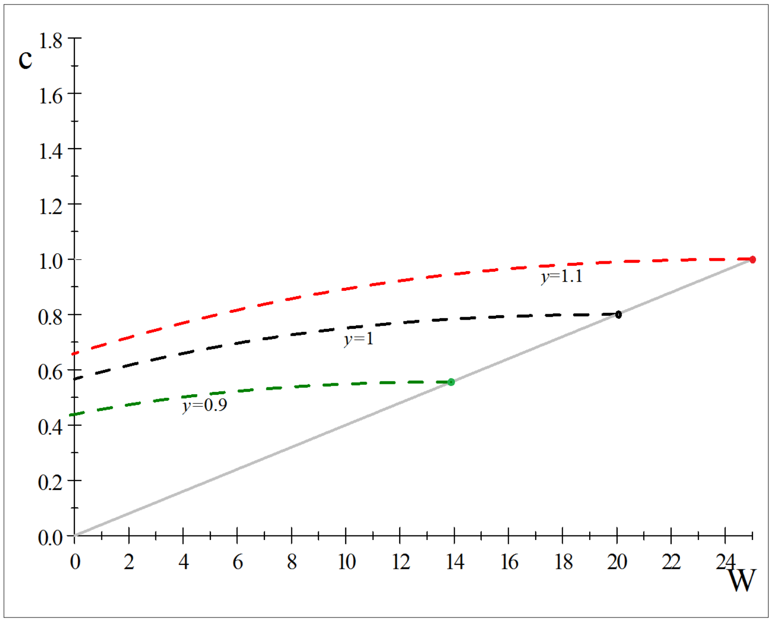
With closed-form solution for the optimal path of saving and consumption, we can provide several additional comparative statics results.

We saw already in Figure 2 how the optimal path of consumption and wealth accumulation depends on the parameter ψ. In (14), higher leisure utility ψ implies a lower retirement threshold W^*. In Figure 2, we see that lower W^* means faster retirement with a higher saving rate along the optimal accumulation path.

We can also examine how consumption, saving, and the retirement decision depend on the level of labor income y. We know from (14) that the retirement threshold wealth level W^* is increasing in y. Using (25), it is not hard to show that if two agents have the same financial wealth W_t and face the same job loss rate λ, the agent with higher labor income y will consume more and retire later. The numerical example given in Figure 5 illustrates this point. In that figure, paths leading to lower retirement points are everywhere below those leading to higher retirement wealth thresholds. Those higher paths correspond to higher labor income y earned during employment.

Finally, we examine how the solution to the agent's optimal consumption, saving, and retirement problem depends on the real interest rate r. Dashed lines in Figure 6 show three accumulation paths, each optimal at a different level of r. That the retirement wealth threshold

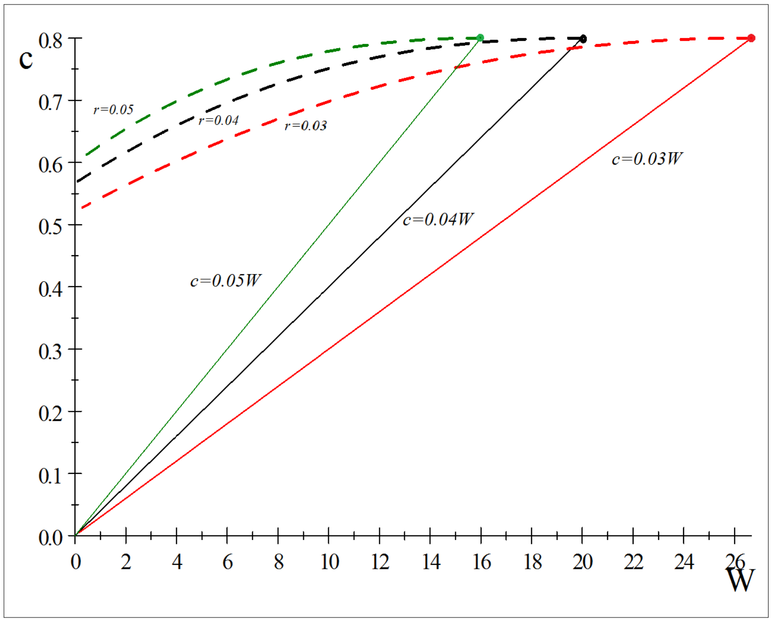
Figure 5 Consumption and Wealth Accumulation for Three Different Values of Labor Income



Notes: Other parameters as in Figure 2.

W^* is lower at higher r can be seen from the terminal points of the accumulation paths in Figure 6, or directly from the formula for W^* given in (14). Since the marginal value of wealth in retirement $u'(rW_t)$ decreases in r, it is intuitive that when r is higher the agent chooses to give up labor income y in return for utility ψ earlier, i.e., at a lower wealth threshold. Figure 6 shows that prior to retirement, at higher r both the agent's consumption c_t and his interest income rW_t are higher. Interest income is represented in Figure 6 by the straight lines rW connecting the origin to the terminal points of the optimal accumulation paths. How the wealth accumulation rate $dW_t/dt = rW_t - c_t + y$ depends on r is determined by the magnitudes of c_t and rW_t. In fact, the rate of wealth accumulation is increasing in r at high levels of wealth W_t and decreasing in r at low levels of W_t. For example, at $W_t = 12$, the vertical distance between (any two) dashed lines (representing c_t) is smaller than the vertical distance between the solid lines

Figure 6 Optimal Consumption and Wealth Accumulation Paths at Three Levels of the Interest Rate



(representing rW_t). At $W_t = 1$, the opposite is true. The cumulative effect of these differences on the agent's wealth is positive. With some algebra that we omit here, it can be shown that the agent retires faster when r is higher. That is, for any given W_t the agent's time to planned retirement, $\tau - t$, is shorter the higher the interest rate r.

8. CONCLUSION

This article studies optimal consumption and saving decisions in an infinite-horizon model that allows for endogenous retirement. Relative to the standard model with no retirement, the optimal saving rate is higher. An increase in the job loss risk decreases consumption, even without the actual job loss occurring. Accounting for retirement subdues the magnitude of the response in consumption to changes in the job loss risk. These results may be important for quantitative analyses of observed consumption and saving decisions.

The strong assumption we make on the shape of the agent's profile of labor income lets us abstract in this article from borrowing constraints. Since his income can only decrease, the agent never wants to borrow in our model, so the no-borrowing constraint is natural in our analysis, and it never binds. Increasing and hump-shaped paths of income are standard in life-cycle models. Incorporating such paths into our model would require an extension of our analysis accounting for the possibility of binding borrowing constraints.

Our analysis of optimal saving for and timing of retirement can be extended to study other types of actions for which savings are important. For instance, due to down-payment requirements, the optimal timing of a house purchase by a household will depend on the financial wealth of the household. Our analysis in this article can be adapted to study jointly the saving decisions and the optimal timing of this purchase.

APPENDIX: APPENDIX A

Proof of Lemma 1

Multiplying (1) by r and subtracting it from (23) we obtain a linear differential equation

$$\frac{d(c_s - rW_s)}{ds} = (r+\lambda)(c_s - rW_s) - ry,$$

which, with the notation $z_s = c_s - rW_s$, we can write more compactly as

$$\frac{dz_s}{ds} = (r+\lambda)z_s - ry. \tag{29}$$

The solution to this equation is standard. Differentiating $z_s e^{-(r+\lambda)s}$, we have

$$d\left(z_s e^{-(r+\lambda)s}\right) = dz_s e^{-(r+\lambda)s} - (r+\lambda)z_s e^{-(r+\lambda)s}ds = -rye^{-(r+\lambda)s}ds,$$

where the second equality uses (29). Integrating from t to τ and solving for z_t yields

$$z_t = \frac{ry}{r+\lambda}\left(1 - e^{-(r+\lambda)(\tau-t)}\right) + z_\tau e^{-(r+\lambda)(\tau-t)}.$$

Writing the boundary condition $c_\tau = rW_\tau + y$ as $z_\tau = y$ and using it in the above general solution gives us (24). With the boundary condition $c_\tau = rW_\tau$, we have $z_\tau = 0$, which gives us (25). For (26), we take a limit of (25) with $\tau \to \infty$.

Proof of Proposition 1

Write (25) as

$$c_t - rW_t = \frac{r}{r+\lambda} y \left(1 - e^{-(r+\lambda)(\tau - t)}\right)$$

and note that the right-hand side of this equality is decreasing in λ and goes to zero as $\lambda \to \infty$. This proves the proposition's conclusions about c_t and, using (1), dW_t/dt. Next, write (27) as

$$W^* - W_t = \frac{\lambda}{r+\lambda} (\tau - t)\, y + \frac{r}{(r+\lambda)^2} \left(1 - e^{-(r+\lambda)(\tau - t)}\right) y$$

and check that the right-hand side of this equality is strictly increasing with respect to both λ and $\tau - t$. Because W^* does not depend on λ, the left-hand side is constant. Thus, the time to retirement $\tau - t$ must decrease when λ increases to keep the right-hand side constant.

APPENDIX: APPENDIX B

Figure 7 provides the analog of Figure 2 for a nonquadratic utility function u. In particular, this figure depicts numerically computed solution paths to the system of differential equations (1)–(22) for constant relative risk aversion (CRRA) preference represented by the utility function u of the form

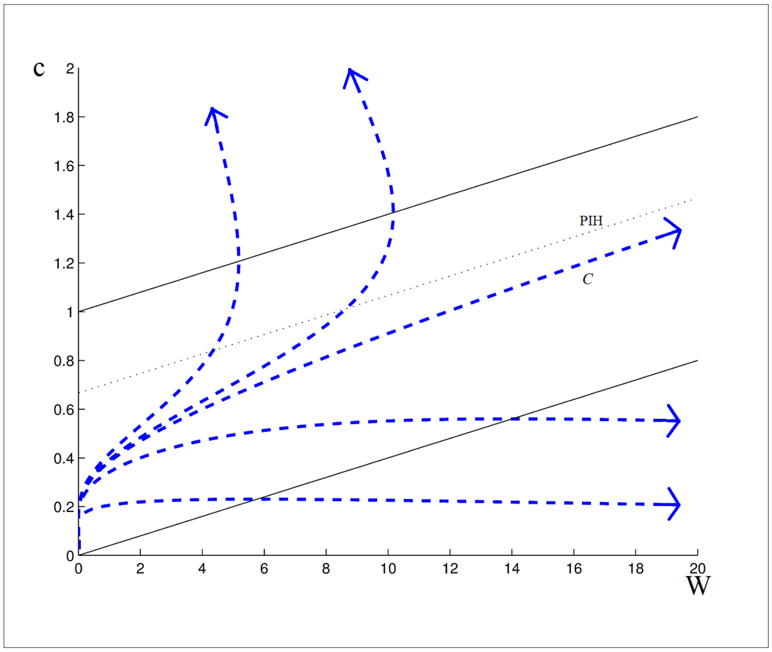
$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}.$$

Qualitatively, these graphs are similar to one another for all values of $\gamma > 0$.

Our analysis determining the optimal accumulation path for a given voluntary retirement wealth threshold W^* from Section 5 is unchanged. The main difference between CRRA preferences and quadratic preferences is that the permanent income hypothesis does not hold under CRRA preferences. With CRRA preferences, agents have the so-called precautionary motive for saving, which is absent under quadratic preferences. When the precautionary saving motive is present, the agent will increase the amount he saves in response to an increase in the riskiness of his income process, holding his expected income constant. (See Ljungqvist and Sargent [2004] for a general discussion of precautionary savings.)

In Figure 7, precautionary savings are best seen by comparing the solution path labelled C with the dotted line labelled PIH. The

Figure 7 Solution Paths for CRRA Preferences



solution path C in Figure 7 is analogous to the solution path C in Figure 2. It is the single solution path that never leaves the middle band of the graph bounded by the lines $c = rW$ and $c = rW + y$. It is the optimal solution path under CRRA preferences for an agent whose $\psi = 0$, i.e., an agent who never chooses to retire voluntarily. The line labelled PIH in Figure 7 is the solution path that would be optimal for that agent if he did not have a precautionary saving motive (i.e., it is an exact replica of the solution path C from Figure 2). At any level of wealth W_t, the vertical distance in Figure 7 between line PIH and the solution path C measures precautionary saving of the agent with CRRA preferences. As we see, precautionary saving is positive at all wealth levels and its magnitude decreases in W_t. In fact, solution C converges to line PIH as $W_t \to \infty$.

As in Figure 2, each solution path crossing the line $c = rW$ is an optimal accumulation path for an agent whose value of the leisure preference parameter ψ is strictly positive. In these cases, as well, precautionary saving can be seen by comparing corresponding solution paths in Figures 2 and 7. For any given voluntary retirement wealth threshold $W^* > 0$, the solution path leading to the retirement point $(W, c) = (W^*, rW^*)$ will in Figure 2 be strictly above the path leading to the same path in Figure 7. The vertical distance between these two paths will represent precautionary saving. All solution paths in Figure 7 converge to zero consumption when wealth goes to zero, while in Figure 2 they do not. Comparing solutions with voluntary retirement in finite time, as in the case of no voluntary retirement, we thus see that the precautionary saving motive is the strongest at very low wealth levels.

APPENDIX: APPENDIX C

In this appendix, we discuss an extension of our model in which the option value of delaying retirement is positive.

Let us add a positive labor income shock to our model. That is, instead of assuming that at all times prior to retirement the agent's labor income is constant, suppose it can increase from y to $\bar{y} > y$. Suppose this upward jump arrives with Poisson intensity $\sigma > 0$. Also, let's assume the job loss shock is independent of the level of income and, as before, it arrives with Poisson intensity $\lambda > 0$.

We will show that with this positive income shock, the agent with income y will not choose to retire as soon as his wealth reaches the threshold W^* but rather will prefer to keep working. The reason why the agent prefers to keep working is that postponing retirement has a positive option value when there is a chance that his labor income increases in the future.

Let $\bar{J}(W_t)$ be the maximal utility value the agent can obtain when his income is already high, i.e., $\bar{y}$. Because once it hits $\bar{y}$ income stays constant until retirement; our previous analysis applies: The agent whose income is $\bar{y}$ will want to retire exactly when his wealth hits the threshold

$$\bar{W}^* = \frac{1}{r} u'^{-1}\left(\frac{\psi}{\bar{y}}\right) > \frac{1}{r} u'^{-1}\left(\frac{\psi}{y}\right) = W^*.$$

We will show, however, that with low income y the agent will not want to retire as soon as his wealth reaches W^*. That is, the retirement rule with wealth threshold level W^* that we obtained in Section 3 is no longer optimal for the agent.

Consider the following strategy for an agent whose labor income is low, y, and whose wealth is W_t. Suppose the agent works over a small time interval $[t, t+h)$. By time $t+h$, three things can happen. The agent loses his job, gets a promotion, or neither. Suppose the agent behaves optimally after a promotion thus obtaining in that event the value $\bar{J}(W_{t+h})$. In the event of the job loss, he behaves optimally in retirement and so he obtains $V(W_{t+h})$. If neither promotion nor job loss happen, suppose the agent retires voluntarily at $t+h$, thus obtaining the value $V(W_{t+h})$ in this event as well. Thus, the agent's strategy is to postpone retirement by h and see if he gets a promotion. If he does not, he quits. Denote by $\bar{V}^h(W_t)$ the value that this strategy gives the agent as of date t.

We proceed analogously to Section 4. We have

$$\begin{aligned}\bar{V}^h(W_t) &= \max_c \Big\{ \textstyle\int_0^h e^{-rs} u(c) ds \\ &\quad + e^{-rh} \Big(e^{-\lambda h}(1 - e^{-\sigma h}) \bar{J}(W_{t+h}) + e^{-\lambda h} e^{-\sigma h} V(W_{t+h}) \\ &\quad + \Big(1 - e^{-\sigma h}\Big) V(W_{t+h}) \Big) \Big\}\end{aligned}$$

with wealth following (1) between t and $t+h$, as the agent works between t and $t+h$. Because h is small, we use the first-order approximation and express $\bar{V}^h(W_t)$ as

$$\begin{aligned}\bar{V}^h(W_t) &= \max_c \big\{ V(W_t) + \big(u(c) + \sigma \bar{J}(W_t) - (r+\sigma) V(W_t) \\ &\quad + V'(W_t) \tfrac{dW_t}{dt} \big) h \big\}.\end{aligned}$$

Next, we compare this value to the value of retiring immediately at t, which we know to be $V(W_t)$. We have that the value of postponing retirement by at least h is strictly preferred to retiring immediately, i.e., $\bar{V}^h(W_t) > V(W_t)$, if and only if

$$\begin{aligned}&\max_c \big\{ V(W_t) + \big(u(c) + \sigma \bar{J}(W_t) - (r+\sigma) V(W_t) \\ &+ V'(W_t)(rW_t + y - c) \big) h \big\} \\ > \; &\max_c \big\{ V(W_t) + \big(u(c) + \psi - rV(W_t) + V'(W_t)(rW_t - c) \big) h \big\}.\end{aligned}$$

Dividing by h, simplifying terms, and removing the identical maximization problems with respect to c on both sides of this condition simplifies it to

$$\sigma \left(\bar{J}(W_t) - V(W_t) \right) + V'(W_t) y > \psi.$$

Now we note that $\bar{J}(W^*) - V(W^*) > 0$ because with high labor income $\bar{y}$ the agent only wants to retire with wealth $\bar{W}^* > W^*$ and not earlier. By definition of W^*, we have $V'(W^*) y = \psi$. Therefore,

$$\sigma \left(\bar{J}(W^*) - V(W^*) \right) + V'(W^*) y > \psi.$$

This means that with low income y and wealth W^*, the agent prefers to postpone retirement. The reason for this is that the term $\sigma\left(\bar{J}(W^*) - V(W^*)\right)$ is strictly positive. This term represents the option value of delaying retirement. For as long as the agent is not retired, he has a chance to see his labor income increase, in which case he would prefer to continue working until his wealth reaches $\bar{W}^*$. Because retirement is permanent, by retiring with wealth $W^* < \bar{W}^*$, the agent closes this possibility to himself or, in other words, gives up this option. By delaying retirement, he keeps this option open.

By continuity, the above condition holds in the neighborhood of W^*, i.e., also for some wealth $W_t > W^*$. At that wealth level we have $V'(W_t)y < \psi$, i.e., in terms of his current payoff the agent would be strictly better off to retire immediately. He does not choose to do so, however, because the option value $\sigma\left(\bar{J}(W_t) - V(W_t)\right)$ is larger than the payoff from retiring $\psi - V'(W_t)y$.

REFERENCES

Bewley, Truman. 1977. "The Permanent Income Hypothesis: A Theoretical Formulation." *Journal of Economic Theory* 16 (December): 252–92.

Chamberlain, Gary, and Charles A. Wilson. 2000. "Optimal Intertemporal Consumption Under Uncertainty." *Review of Economic Dynamics* 3 (July): 365–95.

Farhi, Emmanuel, and Stavros Panageas. 2007. "Saving and Investing for Early Retirement: A Theoretical Analysis." *Journal of Financial Economics* 83 (January): 87–121.

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton, N.J.: Princeton University Press.

Hall, Robert E. 1978. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy* 86 (December): 971–87.

Karatzas, Ioannis, and Steven E. Shreve. 1997. *Brownian Motion and Stochastic Calculus, Second Edition*. New York: Springer Science+Business Media, Inc.

Kingston, Geoffrey H. 2000. "Efficient Timing of Retirement." *Review of Economic Dynamics* 3 (October): 831–40.

Marcet, Albert, Francesc Obiols-Homs, and Philippe Weil. 2007. "Incomplete Markets, Labor Supply and Capital Accumulation." *Journal of Monetary Economics* 54 (November): 2,621–35.

Merton, Robert C. 1971. "Optimum Consumption and Portfolio Rules in a Continuous-Time Model." *Journal of Economic Theory* 3 (December): 373–413.

Leland, Hayne E. 1994. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." *Journal of Finance* 49 (September): 1,213–52.

Ljungqvist, Lars, and Thomas J. Sargent. 2004. *Recursive Macroeconomic Theory, Second Edition*. Cambridge, Mass.: The MIT Press.

Ljungqvist, Lars, and Thomas J. Sargent. Forthcoming. "Career Length: Effects of Curvature of Earnings Profiles, Earnings Shocks, Taxes, and Social Security." *Review of Economic Dynamics*.

Pindyck, Robert S. 1991. "Irreversibility, Uncertainty, and Investment." *Journal of Economic Literature* 29 (September): 1,110–48.