# Accounting for the Non-Employment of U.S. Men, 1968–2010

Marianna Kudlyak, Thomas Lubik, and Jonathan Tompkins

en in their prime working age, defined as men between the ages of 25 and 64, constitute 33 percent of the civilian non-institutionalized population in the United States. At the trough of the 1969–1970 recession, 6.5 percent of this group (henceforth, "population") were out of the labor force (OLF), 90.8 percent were employed, and 2.7 percent were unemployed. Since then, the employment-to-population ratio has trended persistently downward, while the OLF-to-population ratio has increased substantially. In 2010, the aftermath of the 2007–2009 recession, the employment-to-population ratio of this same group declined to an all-time low of 76.3 percent, while the OLF-to-population ratio increased to an all-time high of 14.7 percent (see Figure 1, Panels A–C).

In this article, we investigate the extent to which the change in the sociodemographic composition of the population (by age, educational attainment, marital status, and race) has contributed to the changes in the aggregate labor market outcomes. Our emphasis on the compositional changes in the sociodemographic characteristics of the population is motivated by a literature rife with correlations between sociodemographic factors and labor market outcomes. In particular, older workers typically experience lower rates of labor force participation and, conditional on participating, older workers are

We would like to thank Huberto Ennis, Arantxa Jarque, Nadezhda Malysheva, and Alexander Wolman for their invaluable comments. The views expressed here do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: marianna.kudlyak@rich.frb.org; thomas.lubik@rich.frb.org; jonathan.tompkins@rich.frb.org.

<sup>&</sup>lt;sup>1</sup> The focus of this article is on the employment-, unemployment- and OLF-to-population ratios. These are defined as the proportion of individuals in the entire population with a given labor status. They are thus distinct from rates (e.g., the unemployment rate), which are defined as the proportion of the labor force (i.e., the sum of unemployed and employed persons) with a given labor status outcome.

Panel A: Employment-to-Population Proportion of Population 1970 1980 1990 2000 Panel B: Unemployment-to-Population .10 Proportion of Population 1970 1980 2000 2010 1990 Year Panel C: OLF-to-Population .16 Proportion of Population 1980 1990 2000 2010

Figure 1 Labor Market Outcomes of 25-64-Year-Old Men

Notes: Authors' own calculations from the IPUMS-CPS data. Shaded areas on this and subsequent graphs represent National Bureau of Economic Research (NBER)-dated recessions.

less likely to be unemployed than younger workers (see, for example, Shimer 1999). The literature also finds that (i) more highly educated workers have a higher opportunity cost of not working; (ii) married men are more likely to participate in the labor force and, conditional on participation, more likely

to be employed; and (iii) non-white persons are usually underrepresented in the labor force and employment. Thus, one expects a strong association between labor market outcomes and the demographic composition of the labor force, which serves as a reduced-form representation of underlying structural relationships.

In this article, we decompose the observed changes in aggregate labor market outcomes into changes in the sociodemographic composition of the population and changes in the labor market outcomes of different sociodemographic groups. For each year we generate two sets of counterfactual aggregate labor market outcomes. The first set is generated by using the sociodemographic composition of the population from all the years in the sample and holding the labor market outcomes of different sociodemographic groups constant at the actual level of the reference year. The second set is generated by holding the sociodemographic composition constant instead. We then use these counterfactuals to perform the decomposition of the changes in the aggregate labor market outcomes. Finally, we use the most recent sociodemographic composition of the population to forecast the aggregate OLF-to-population ratio in 2015.

Given the similarities between the 1980–1982 and 2007–2009 recessions in terms of severity, we emphasize, throughout this article, comparisons between the labor market outcomes in 1983 and 2010. We find that the changes in the demographic composition of the population explain much of the historical upward trend in the OLF ratio. The OLF ratio increased from 11.1 percent in 1983 to 14.7 percent in 2010. Of this increase, 1.9 percentage points (49 percent of the total change) are attributable to changes in the sociodemographic composition of the population. The employment-to-population ratio fell from 80.2 percent in 1983 to 76.3 percent in 2010. We find that changes in the sociodemographic composition of the population account for 1.7 percentage points (44 percent) of this decline. The unemployment-to-population ratio increased from 8.7 percent in 1983 to 8.9 percent in 2010, but none of this increase can be accounted for by changes in the demographic composition of the population. Finally, using predicted changes in the age distribution of the population, we estimate that the OLF-to-population ratio will increase to more than 16 percent in 2015 as compared to 14.7 percent in 2010.

When interpreting our results we need to be wary that changes in the sociodemographic composition might cause changes in the labor market outcomes of different sociodemographic groups. Alternatively, changes in the labor market outcomes of some sociodemographic groups might cause changes in the sociodemographic composition of the population. For example, an increase in the employment probability for higher educated workers relative to other education levels might contribute to an increase in the educational attainment of the population. Alternatively, an increase in educational attainment of

the population can change the employment probabilities of different groups. Our accounting exercise does not account for these effects.

This article is related to the existing literature that documents a secular decline in the labor force participation of prime working age men. Autor and Duggan (2003) document a substantial fall in labor force participation among men. Using data from the Current Population Survey (CPS), Juhn, Murphy, and Topel (1991, 2002) find that falling unemployment rates among men in the 1990s greatly exaggerated the improvements of the labor outcomes for this population because the period was also characterized by a fall in the labor force participation rate. We update their analyses by focusing on the decomposition of the changes in the labor outcomes into changes in the sociodemographic characteristics of the population and by adding data from the most recent decade. Our work is also closely related to Little and Bradley (2007), who use a multinomial logistic model to study the sociodemographic determinants of labor market outcomes, distinguishing employment, unemployment, peripheral inactivity (marginally attached to labor force), and OLF, in Great Britain. Finally, our work expands on that of Fallick and Pingle (2006), who decompose movements in U.S. labor force participation into aging of the population and labor force trends within age groups. Whereas these authors focus solely on changes in the OLF-to-population ratio caused by aging of the population. we consider changes in each employment status caused by changes in four different sociodemographic factors.

The article is structured as follows. Section 1 describes the data. Section 2 summarizes the changes in the demographic composition of the population of working age men between 1968–2010 and documents trends in labor market outcomes by sociodemographic groups. Section 3 describes the decomposition exercise and presents results of the decomposition of changes in labor outcomes between 2010 and the earlier years. Section 4 presents the forecast of the 2015 OLF-to-population ratio. Section 5 concludes.

#### 1. DATA

We use data from the Integrated Public Use Microdata Series CPS (IPUMS-CPS), which comprises data from the March Supplement of the Current Population Survey (hereafter referred to as the March CPS). The CPS is a monthly survey of U.S. households' activities, conducted by the Census Bureau and the Bureau of Labor Statistics and designed to measure unemployment. The basic survey is conducted every month; over time various supplementary surveys have been conducted to study different social and economic questions. The March CPS contains in-depth information on sociodemographic characteristics of the population and income. The variables in IPUMS-CPS are coded identically or "harmonized" over the years.

The CPS is a collection of individual-level data obtained from the interviewed households. We focus on males between the ages of 25 and 64.<sup>2</sup> Throughout the analysis we use the March CPS sampling weights that account for a complex survey design. The aggregate annual statistics that we report thus correspond to March of a respective year.

It should be noted that in 1994 the CPS underwent a major redesign both in the wording of its questions and in the methodology of the data collection process, which led to some discrepancies between the aggregate series constructed from the microdata prior to the redesign and post-redesign. However, it is not a concern of our analysis because the inconsistencies associated with the aggregate labor statistics for the sample of 25–64-year-old men are minor and not statistically significant (see Polivka and Miller [1995]).

### 2. SOCIODEMOGRAPHIC COMPOSITION OF THE POPULATION AND LABOR OUTCOMES BY DIFFERENT GROUPS

Between 1968–2010 there have been considerable changes in the distribution of 25–64-year-old civilian, non-institutionalized men by age and education, and some noticeable changes by marital status and race (Figure 2). Figures 3–5 display the time series of labor outcomes by different sociodemographic groups. In general, across different groups the employment-to-population ratio has been trending down, while the OLF-to-population has been gradually increasing. We now describe each figure in detail.

#### Sociodemographic Composition

Panel A of Figure 2 shows the changes in the shares of 25–34, 35–44, 45–54, and 55–64-year-old men in the population. From 1968–1986, the share of 25–34-year-old men grew steadily, reaching its largest fraction of 35.6 percent in 1986, and declined thereafter. The share of 55–64-year-old men fell from 1968–1995, reaching its smallest fraction of 15 percent in 1995, and has increased steadily since. From 2000 to present, the shares of older workers (45–54 and 55–64-year-olds) have been increasing, while the shares of younger workers (25–34 and 35–44-year-olds) have been decreasing. This shift in the age distribution toward older workers is largely because of the

 $<sup>^2</sup>$  This article focuses on the male population. The OLF-to-population ratio for women fell drastically from 1968 to the mid-1990s as females entered employment, while that of males trended upward. Since the mid-1990s, the OLF-to-population ratios for males and females have experienced similar trends, though the OLF-to-population ratio for females remains approximately 10 percentage points higher. While studying aggregate employment outcomes for both genders would be an interesting exercise, doing so is beyond the scope of this article.

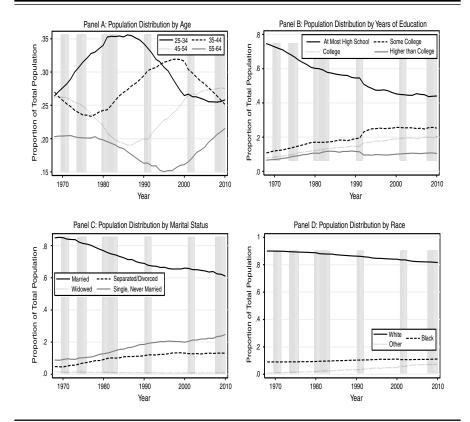


Figure 2 Distribution by Age, Education, Marital Status, and Race

Notes: Authors' own calculations from the IPUMS-CPS data. In this graph, "population" refers to the population of 25–64-year-old civilian, non-institutionalized men.

aging of the Baby Boom generation. In 2010, the shares of 25–34, 35–44, 45–54, and 55–64-year-olds were 25.8, 25.2, 27.5, and 21.5, respectively.

Panel B of Figure 2 shows the upward trend in the educational attainment of the population and reveals that the shares of men with some college, college, or higher than college education have been increasing at the expense of men with at most a high school degree. The share of the latter has declined from 74.6 percent in 1969 to 44.0 percent in 2010. Panel C shows that the population distribution by marital status has shifted toward divorced or separated and single men at the expense of married men. In 1968, 84.8 percent of the 25–64-year-old men were married, while only 61.1 percent were married in 2010. Finally, Panel D shows that the share of white men in the population has been

Panel A: Proportion of Males Employed by Age Panel B: Proportion of Males Employed by Years of Education 35-44 55-64 .95 45-54 .90 .85 .80 .75 At Most High School --- Some College Higher than Colle College 1990 2000 1970 1980 2010 1970 1990 2000 Year Panel C: Proportion of Males Employed by Marital Status Panel D: Proportion of Males Employed by Race Single, Never Married 1990 2000 2010 1970 1990 2000 2010 Year Year

Figure 3 Employment-to-Population Ratio by Sociodemographic Group, 25–64-Year-Old Males

Notes: Authors' own calculations from the IPUMS-CPS data.

steadily falling over the last 40 years while the shares of black men and men of other races have been increasing.

#### **Employment-to-Population Ratio**

Employment-to-population ratios by age are shown in Figure 3, Panel A. The figure shows that the proportion of employed workers for the 25–34, 35–44, and 45–54 age groups move in sync over the years, declining from roughly .95 in 1968 to approximately .80 in 2010. The 55–64 age group has displayed markedly different behavior, declining from approximately .82 in 1968 to an all-time low of .60 in 1994 before trending back up to .65 in 2010.

Figure 3, Panel B displays the employment-to-population ratio by educational attainment. For each group there is a clear decline in the employment-to-population ratio over time, with the severity of this decline decreasing in years of schooling. From 1968–2010, those with more than a college education and those with just a college degree have seen moderate declines of .06 and .11, respectively, while those with some college experience and those with at most a high school degree have experienced larger declines of .18 and .21, respectively.

The employment-to-population ratio trends by marital status are displayed in Figure 3, Panel C. Though there have been decreases in the employment-to-population ratio across marital status groups from 1968–2010, we observe that much of these declines came prior to 1980 and after 2007. Between these two dates, the employment-to-population ratio across marital status groups shows little or no trend.

Panel D of Figure 3 shows the employment-to-population ratio for each race. We observe declines in the employment-to-population ratio for whites, blacks, and others over the sample period, with this decline being most pronounced for blacks. The employment-to-population ratio fell by approximately .24 for this group from 1968–2010, whereas the employment-to-population ratio for whites and others declined by .14 and .10, respectively.

#### **Unemployment-to-Population Ratio**

The unemployment-to-population ratio shows a clear cyclical pattern, rising during economic contractions and falling during expansions. As can be seen from Panel A of Figure 4, the unemployment-to-population ratio is decreasing with age. The rise in the unemployment-to-population ratio across age groups from 2007–2010 is comparable to the increase from 1980–1983, though the increase for the 25–34 age group was more pronounced in the 1980–1983 recession, and the increases for the 35–44 and 45–54 age groups are more pronounced in the 2007–2009 recession.

The unemployment-to-population ratios for educational attainment groups are displayed in Figure 4, Panel B, which shows that more years of schooling are associated with lower unemployment. Figure 4, Panel C shows the unemployment-to-population ratio by marital status and shows that single, never married and separated/divorced individuals have consistently higher unemployment-to-population ratios than those who are married. Finally, as can be seen from Panel D of Figure 4, the unemployment-to-population ratio for blacks is consistently higher than the unemployment-to-population ratio for whites and others. Between 2007 and 2010, the unemployment-to-population ratio for blacks increased by .08, whereas the ratios of whites and others increased by approximately .04.

Panel A: Proportion of Males Unemployed by Age Panel B: Proportion of Males Unemployed by Years of Education .12 At Most High School ---- Some College .10 N۶ .06 .04 02 2000 1970 1980 1990 2000 1990 Year Year Panel C: Proportion of Males Unemployed by Marital Status Panel D: Proportion of Males Unemployed by Race .15 Single, Never Marrie Other .10 1980 1990 2000 2010 Year Year

Figure 4 Unemployment-to-Population Ratio by Sociodemographic Group, 25–64-Year-Old Males

Notes: Authors' own calculations from the IPUMS-CPS data.

#### **OLF-to-Population Ratio**

Figure 5, Panel A displays the OLF-to-population ratio for each age group. We note that each time series has a distinct upward trend from 1968–2010, with the upward trend for the 55–64 age group from 1968–1985 being particularly severe. However, the OLF-to-population ratio for the 35–44 age group has been relatively stable since 1994 and that of the 55–64 age group has actually declined moderately since 1994. The most notable feature of this figure is the disparity between the 55–64 age group and the other age groups. Historically, the OLF-to-population ratio of the 55–64 age group has dwarfed that of the younger age groups. In 2010, the OLF-to-population ratio for the 55–64 age group was .30, whereas the ratios of the other age groups ranged from .09

Panel A: Proportion of Males OLF by Age Panel B: Proportion of Males OLF by Years of Education At Most High School ---- Some College .15 25-34 35-4 55-64 .10 .05 1970 1980 1990 2000 2010 1970 1980 1990 2000 2010 Panel C: Proportion of Males OLF by Marital Status Panel D: Proportion of Males OLF by Race .25 .20 Single, Never Married .15 1970 1980 1990 2000 2010 1970 1980 1990 2000 2010 Year

Figure 5 OLF-to-Population Ratio by Sociodemographic Group, 25–64-Year-Old Males

Notes: Authors' own calculations from the IPUMS-CPS data.

to .13. We also observe that the OLF-to-population ratio of the 45–54 group tends to be higher than the 25–34 and 35–44 age groups by approximately .02.

The OLF-to-population ratios for different educational attainment groups are shown in Panel B of Figure 5. There are upward trends since 1968 for each group, with the largest increases occurring for those with at most a high school education and those with some college education. The figure shows that those with a college degree and those with more than a college education have experienced similar OLF-to-population ratios across time, though since 2005 these series have diverged, with the OLF-to-population ratio for those with a college degree continuing to trend upward while the ratio for those with more than a college education has fallen. Finally, we observe that fewer years of education are associated with a larger OLF-to-population ratio: In 2010,

the OLF-to-population ratios were .19, .15, .09, and .07 for those with at most a high school degree, some college education, a college degree, and more than a college degree, respectively.

Panel C of Figure 5 gives the OLF-to-population ratio by marital status. There is a notable upward trend in the OLF-to-population ratio for each marital status group. Those in the married group have the lowest OLF-to-population ratio throughout the observation period, while the separated/divorced and single, never married groups have an OLF-to-population ratio that is between .05 and .1 higher. The widowed group has had the highest OLF-to-population ratio historically, occasionally exceeding .40.

Figure 5, Panel D breaks down the OLF-to-population ratio by race. The most notable feature of this figure is the large difference in the growth of the OLF-to-population ratio between blacks and the other groups. The OLF-to-population ratio for blacks has increased from .09 in 1968 to .24 in 2010, whereas the OLF-to-population ratios for the other and white groups have increased from .09 and .06 to .15 and .14, respectively. Thus, while each series has trended upward, that of blacks has done so more rapidly.

#### 3. ACCOUNTING FOR CHANGES IN AGGREGATE LABOR OUTCOMES

#### Method

In this subsection we discuss the methods by which we create counterfactual labor outcomes and decompose the changes in actual labor market outcomes.

The aggregate share of persons with labor outcome LO, where  $LO = \{\text{Employed, Unemployed, OLF}\}$ , in year t can be described by the following equation:

$$\frac{LO_t}{pop_t} = \sum_{i} \left( \frac{LO_{i,t}}{g_{i,t}} \times \frac{g_{i,t}}{\sum_{i} g_{i,t}} \right), \tag{1}$$

where  $i \in A \times E \times M \times R$  corresponds to a vector of demographic characteristics consisting of age (A), educational attainment (E), marital status (M), and race (R);  $g_{i,t}$  is the number of persons in group i;  $LO_{i,t}$  is the number of persons with the labor status LO in group i; and  $pop_t$  is the size of the population.

In essence, we divide the population into mutually exclusive groups (e.g., married, college-educated white males between the ages of 25-34). Equation (1) describes the aggregate proportion as the sum of the labor outcomes by group (the first term in the equation) weighted by the size of the groups in the population (the second term in the equation). For example, fixing either term for all i at its 2010 level while allowing the other to take on historical values allows us to construct counterfactual aggregate labor outcomes for 2010. By

creating and comparing time series of these counterfactuals, we can observe the degree to which these two terms are driving changes in aggregate labor outcomes in 2010.

One concern when creating these counterfactuals is that the changing sociodemographic composition of the population could affect the labor outcomes of different groups. Alternatively, changes in the labor outcomes of a sociodemographic group could change the sociodemographic composition of the population. For example, if college-educated individuals are more likely to be employed relative to other educational attainment subgroups, there will likely be an influx of individuals into the college-educated demographic group. By simply fixing either of these terms while varying the other, we do not account for these endogeneity effects.

To analyze the change in labor outcomes, we perform the following decomposition:

$$\frac{LO_{t_2}}{pop_{t_2}} - \frac{LO_{t_1}}{pop_{t_1}} = \left[\sum_{i} s_{i,t_2,t_1} - \sum_{i} s_{i,t_1,t_1}\right] + \left[\sum_{i} s_{i,t_2,t_2} - \sum_{i} s_{i,t_2,t_1}\right], (2)$$

where 
$$s_{i,t_x,t_y} = \frac{LO_{i,t_x}}{g_{i,t_x}} \times \frac{g_{i,t_y}}{\sum g_{i,t_y}}$$
.

The component in the first set of brackets in equation (2) measures the effect of changes in the labor outcomes of different groups from year  $t_1$  to  $t_2$ , given the sociodemographic composition of the population in year  $t_1$ . The second term captures the effect of changes in the sociodemographic composition of the population, given the labor outcomes of different groups in year  $t_2$ .

Alternatively, we can write

$$\frac{LO_{t_2}}{pop_{t_2}} - \frac{LO_{t_1}}{pop_{t_1}} = \left[\sum_{i} s_{i,t_1,t_2} - \sum_{i} s_{i,t_1,t_1}\right] + \left[\sum_{i} s_{i,t_2,t_2} - \sum_{i} s_{i,t_1,t_2}\right].$$
(3)

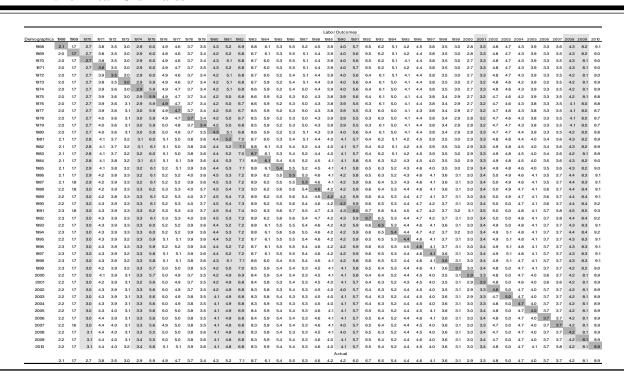
The component in the first set of brackets in equation (3) measures the effect of changes in the sociodemographic composition of the population, given the labor outcomes of different groups in year  $t_1$ . The second term captures the effect of changes in the labor outcomes of different groups from year  $t_1$  to  $t_2$ , given the sociodemographic composition of the population in year  $t_2$ . The difference between these two decompositions is the base year, i.e., the year at which the component, other than the component of interest, is held constant. For example, the change as a result of a change in the labor outcomes of different groups in equation (2) is calculated using  $t_1$  as the base year, while in equation (3) it is calculated using  $t_2$  as the base year. Because of the endogeneity issues mentioned above, these two decompositions do not necessarily deliver the same results. It is also unclear that one is theoretically better than the other. We perform both decompositions and, despite some

Table 1 Counterfactual Predictions of the Employment-to-Population Ratio by Percent, 25–64-Year-Old Males

																																															_
																									Labor																						
emographic																																															
1968																																											82.5				
1969																																											82.6				
1970																																											82.6				
1971																																											82.6				
1972						_																																					82.7				
1973							_																																				82.9				
1974		915																																									82.8				
1975	91.6	510																																									82.9				
1976	91.7									_	_																																83.1				
1977																																											83.1				
1978																																											1 83.1 2 83.2				
1979																																											83.2				
1981																																											83.2 83.3				
1982																_																											83.4				
1982																																											83.4				
1984																																											83.6				
1985		913																		_																							83.6				
1986		913																			_																						83.6				
1987		914																				_																					7 83.8				
1988	914	913																																									83.6				
1989																																											83.7				
1990	914	913	90.0	88.3	88.0	8 0	8.2	88.2	84.5	84.	.7 8	85.3	86.3	86.9	85.7	84.4	82.	4 80	2 82	.8 8:	3.3 8	3.0	83.5	84.1	84.6	84.5	82.8	819	817	815	82.7	82.9	83.6	84.4	84.5	84.7	84.2	82.6	82.1	82.2	83.0	83.7	7 83.8	82.	.8 78.	3 78	3.0
1991	913	912	89.9	88.2	87.5	9 8	8.0	0.88	84.5	84.	.6	85.1	86.2	86.7	85.6	84.2	82.	2 80	0 82	.6 8	3.1 8	2.8	83.3	83.8	84.3	84.3	82.6	81.7	815	813	82.5	82.6	83.4	84.2	84.3	84.5	84.0	82.4	81.9	819	82.8	83.5	83.6	82.	.6 78	.1 77	8.5
1992	912	912	89.9	88.2	87.5	9 8	38.1	0.88	84.6	84.	.6 8	85.2	86.3	86.9	85.7	84.4	82.	4 80	2 82	.8 8:	3.2 8	3.0	83.5	84.0	84.6	84.5	82.9	819	816	815	82.7	82.8	83.6	84.3	84.5	84.7	84.2	82.6	82.1	82.1	82.9	83.6	83.8	82.	8 78.	3 77	.9
1993	912	912	89.9	88.2	87.5	8 6	8.2	0.88	84.7	84.	.7 8	85.2	86.3	86.9	85.8	84.4	82.	5 80	3 82	9 8:	3.3 8	3.0	83.5	84.0	84.6	84.5	82.9	82.0	816	815	82.7	82.8	83.6	84.3	84.5	84.7	84.2	82.6	82.1	82.1	82.9	83.6	83.9	82.	9 78.	4 78	J.0
1994	912	912	89.9	88.	87.9	9 8	38.1	0.88	84.8	84.	.8 8	85.3	86.4	87.0	85.9	84.5	82.	5 80	3 83	.0 8:	3.3 8	83.1	83.5	84.1	84.7	84.6	83.0	82.1	817	816	82.8	82.8	83.7	84.4	84.6	84.8	84.3	82.7	82.1	82.1	83.0	83.7	83.9	83.	.0 78.	5 78	8.1
1995	912	913	90.0	88.2	88.0	8 0	8.3	0.88	85.0	85.	.o s	85.5	86.6	87.2	86.0	84.6	82.	7 80	5 83	3.1 83	3.5 8	3.2	83.7	84.2	84.8	84.8	83.2	82.3	819	818	83.0	83.0	83.9	84.5	84.8	84.9	84.5	82.9	82.3	82.3	83.2	83.9	84.1	1 83.	2 78.	7 78	1.3
1996	91	911	89.8	88.0	87.5	8 6	8.0	87.9	84.7	84.	.8 8	85.3	86.4	86.9	85.8	84.5	82.	6 80	4 83	.0 8:	3.3	83.1	83.5	84.0	84.7	84.6	83.1	82.1	817	816	82.7	82.8	83.7	84.3	84.6	84.7	84.3	82.7	82.1	82.1	83.0	83.7	83.9	83.	.0 78.	5 78	3.1
1997	90.9	91	89.7	87.8	87.	8	7.9	87.8	84.6	84.	.7	85.1	86.3	86.8	85.7	84.4	82.	5 80	3 82	9 8	3.2 8	3.0	83.4	83.9	84.6	84.5	83.0	82.0	815	815	82.6	82.7	83.6	84.3	84.4	84.6	84.2	82.6	81.9	819	82.9	83.6	83.8	82.	9 78.	4 78	1.0
1998	90.9	911	89.7	87.8	87.	7 8	7.9	87.6	84.6	84.	.6 8	85.2	86.3	86.9	85.7	84.4	82.	5 80	3 82	9 8	3.2 8	3.0	83.3	83.8	84.5	84.4	82.9	819	815	814	82.5	82.6	83.5	84.2	84.4	84.5	84.1	82.5	81.9	819	82.8	83.4	83.7	82.	9 78.	4 78	1.0
1999																																											83.7				
2000																																											83.8				
2001																																											83.6				
2002																																											83.4				
2003																																											83.2				
2004																																											83.0				
2005																																										_	82.8				
2006																																											82.6				
2007																																											82.5	_	in a		
2008																																											82.3			_	
2009																																											82.1				
2010	89.1	89.4	87.8	86.	85.9	9 8	6.2	85.8	82.9	82.	.5 8	83.3	84.6	85.1	84.	82.5	80.	9 78	5 80	8 8	12 8	9.00	81.4	817			810	80.1	79.4	79.3	80.4	80.3	814	82.1	82.4	82.4	82.1	80.5	79.9	80.0	810	815	818	81	11 76.	7 76	3
																										Actual																	82.5				

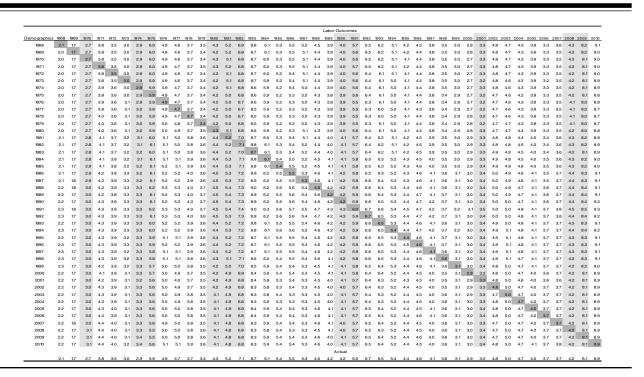
Notes: Authors' own calculations from the IPUMS-CPS data. The columns correspond to the year of the labor outcomes used. The rows correspond to the year of the sociodemographic composition used. Lightly shaded columns correspond to the NBER-dated contractions (from peak to trough). Darkly shaded elements correspond to actual values for given year.

Table 2 Counterfactual Predictions of the Unemployment-to-Population Ratio by Percent, 25–64-Year-Old Males



Notes: Authors' own calculations from the IPUMS-CPS data. The columns correspond to the year of the labor outcomes used. The rows correspond to the year of the sociodemographic composition used. Lightly shaded columns correspond to the NBER-dated contractions (from peak to trough). Darkly shaded elements correspond to actual values for given year.

Table 3 Counterfactual Predictions of the OLF-to-Population Ratio by Percent, 25–64-Year-Old Males



Notes: Authors' own calculations from the IPUMS-CPS data. The columns correspond to the year of the labor outcomes used. The rows correspond to the year of the sociodemographic composition used. Lightly shaded columns correspond to the NBER-dated contractions (from peak to trough). Darkly shaded elements correspond to actual values for given year.

small quantitative differences, find the qualitative conclusions from the two decompositions to be the same. For this reason, we report only the results of the decomposition corresponding to equation (2).<sup>3</sup>

#### **Changes in Aggregate Labor Outcomes**

Tables 1–3 show the predicted labor outcomes—proportion employed, unemployed, and OLF, respectively—calculated using the labor outcomes of different groups from year  $t_1$  and the demographic composition of the population from year  $t_2$ , where  $t_1, t_2 \in [1968, 2010]$ .

The diagonal elements in Tables 1–3 (darkly shaded) show the actual labor outcomes for their respective year. The off-diagonal elements show the counterfactual labor outcomes. Thus, for each year we have two sets of counterfactual predictions. Moving along a row gives the predicted labor outcome for a fixed demographic composition, while moving down a column gives the predicted labor outcome for fixed labor outcomes of different groups. For example, the (1983, 2010) entry of Table 1 gives the predicted proportion of employed individuals in 2010 given the 1983 demographic composition (77.9 percent), while the (2010, 1983) entry gives the predicted proportion of employed individuals in 2010 given the 1983 labor outcomes of different groups (78.6 percent).

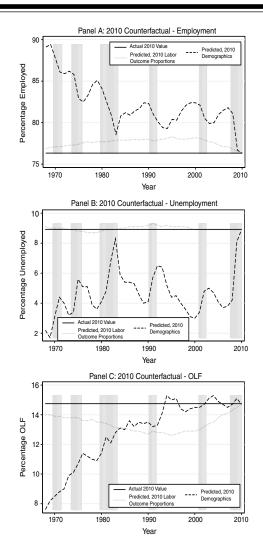
#### **Employment**

The actual 2010 value of the employment-to-population ratio for 25–64-year-old men and the two series of counterfactual predictions for 2010 are shown in Figure 6, Panel A. The dashed line shows the predicted employment-to-population ratio from holding the demographic composition of the population constant at its 2010 level but varying the labor outcomes of different groups. The dotted line shows the predicted employment-to-population ratio from holding the labor outcomes of different groups at their 2010 level but varying the sociodemographic composition of the population. The actual employment-to-population ratio in 2010 is lower than any point of the predicted counterfactual series. This implies that changes in both the sociodemographic composition and labor outcomes of different groups in 2010 contribute to the historically low employment-to-population ratio among men.

To examine the contribution of the change in the labor outcomes of different groups and the sociodemographic composition to changes in the employment-to-population ratio, we construct each term of equation (2) for  $t_1 \in [1968, 2010]$  and  $t_2 = 2010$ . Table 4 reports the change in the employment-to-population ratio, the total change as a result of change in each

<sup>&</sup>lt;sup>3</sup> Results of the alternative decomposition are available upon request.

Figure 6 Counterfactual Predictions of Labor Outcomes, 25–64-Year-Old Males



Notes: Authors' own calculations from the IPUMS-CPS data. The figure shows the outcome-to-population ratio for each of the three labor outcomes. "Predicted, 2010 Demographics" gives the predicted 2010 labor outcome using the 2010 demographic composition and historical labor outcome proportions. "Predicted, 2010 Labor Outcome Proportions" gives the predicted 2010 labor outcome using the 2010 labor outcome proportions and historical demographic compositions.

of the two terms, and the percentage of the total change that is accounted for by changes in each of the two terms. Figure 7, Panel A plots the change in the employment-to-population ratio and the total change as a result of each of the two terms. We find that the change in the labor outcomes of different groups accounts for the majority of the change in the employment-to-population ratio. Comparing the aftermath of the 2007–2009 and 1980–1982 recessions, we see that between 1983–2010 the employment-to-population ratio fell by 3.86 percentage points, of which 41.7 percent is a result of the change in the demographic composition.

We conclude that the decline in the employment-to-population ratio is a result of both changes in the sociodemographic composition and changes in the labor outcomes of different groups.

#### Unemployment

Panel B of Figure 6 plots the two counterfactual series of the 2010 unemployment-to-population ratio against its actual 2010 value. We draw two key observations from the figure: (1) the actual 2010 unemployment-to-population ratio is higher than any point of the predicted counterfactual series that holds the sociodemographic composition constant at its 2010 level; and (2) the actual 2010 unemployment-to-population ratio is lower than the predicted counterfactual for some periods when we hold the labor outcomes of different groups constant at their 2010 level.

These observations suggest that (1) the labor outcomes of different sociodemographic groups in 2010 contribute to a higher unemployment-to-population ratio than the labor outcomes of different groups in all previous years; and (2) the sociodemographic composition of the population in 2010 actually contributes to a lower unemployment-to-population ratio as compared to the sociodemographic composition in some earlier years.

Table 5 breaks down the total change in the unemployment-to-population ratio between a given year and 2010 into change caused by developments in the demographic composition of the population and change caused by developments in the labor outcomes of different groups. Figure 7, Panel B plots the total change in the unemployment-to-population ratio and the change as a result of each of the two terms. The results of this table and graph corroborate our above claims. Changes in the demographic composition of the population have contributed a small, and often negative, amount of the increase in the unemployment-to-population ratio, whereas changes in the labor outcomes of different groups have been responsible for approximately 100 percent of the increase. We observe that, relative to 1983, the 2010 unemployment-to-population ratio is 0.18 percentage points higher, of which 101.0 percent of the change is a result of a change in the labor outcomes of different groups. Thus, the rise in the 2010 unemployment-to-population ratio relative to its

**Table 4 Decomposition of the Change in the Employment-to-Population Ratio Between 2010 and Earlier** Years, 25–64-Year-Old Males

Year	Total Change	9	Because of Outcomes	Change Because of Sociodemographics					
		Percentage Points	% of Total Change	Percentage Points	% of Total Change				
1968	-15.52	-14.96	96.39	-0.56	3.61				
1969	-15.52	-14.87	95.81	-0.65	4.19				
1970	-14.45	-13.70	94.78	-0.76	5.22				
1971	-12.75	-12.03	94.32	-0.72	5.68				
1972	-12.46	-11.61	93.18	-0.85	6.82				
1973	-12.40	-11.40	91.92	-1.00	8.08				
1974	-12.22	-11.26	92.11	-0.96	7.89				
1975	-8.26	-7.23	87.47	-1.04	12.53				
1976	-8.61	-7.33	85.20	-1.27	14.80				
1977	-9.08	-7.75	85.36	-1.33	14.64				
1978	-9.79	-8.54	87.21	-1.25	12.79				
1979	-10.30	-8.90	86.41	-1.40	13.59				
1980	-9.19	<b>-7.77</b>	84.59	-1.42	15.41				
1981	-7.91	<b>−6.57</b>	83.10	-1.34	16.90				
1982	-5.83	<b>-4.37</b>	74.85	-1.47	25.15				
1983	-3.86	-2.25	58.29	-1.61	41.71				
1984	-6.25	-4.65	74.44	-1.60	25.56				
1985	-6.81	-5.25	77.03	-1.56	22.97				
1986	-6.56	-5.03	76.73	-1.53	23.27				
1987	-7.17	-5.50	76.69	-1.67	23.31				
1988	-7.66	-6.07	79.23	-1.59	20.77				
1989	-8.26	-6.59	79.68	-1.68	20.32				

Table 4 (Continued) Decomposition of the Change in the Employment-to-Population Ratio Between 2010 and Earlier Years, 25–64-Year-Old Males

Year	Total Change	· ·	Because of Outcomes	Change Because of Sociodemographics					
		Percentage Points	% of Total Change	Percentage Points	% of Total Change				
1990	-8.20	-6.49	79.15	-1.71	20.85				
1991	-6.28	-4.80	76.45	-1.48	23.55				
1992	-5.61	-4.01	71.44	-1.60	28.56				
1993	-5.32	-3.61	67.81	-1.71	32.19				
1994	-5.28	-3.53	66.78	-1.76	33.22				
1995	-6.62	-4.61	69.59	-2.01	30.41				
1996	-6.48	-4.69	72.39	-1.79	27.61				
1997	-7.25	-5.56	76.65	-1.69	23.35				
1998	-7.82	-6.17	78.87	-1.65	21.13				
1999	-8.04	-6.29	78.27	-1.75	21.73				
2000	-8.22	-6.38	77.57	-1.84	22.43				
2001	-7.62	-5.90	77.39	-1.72	22.61				
2002	-5.89	-4.33	73.61	-1.55	26.39				
2003	-5.07	-3.66	72.14	-1.41	27.86				
2004	-4.92	-3.69	75.03	-1.23	24.97				
2005	-5.59	-4.62	82.75	-0.96	17.25				
2006	-6.00	-5.24	87.39	-0.76	12.61				
2007	-6.20	-5.46	88.10	-0.74	11.90				
2008	-5.28	-4.80	91.00	-0.47	9.00				
2009	-0.68	-0.39	56.79	-0.29	43.21				
2010	0.00	0.00	0.00	0.00	0.00				

Notes: Authors' own calculations from the IPUMS-CPS data. Bold rows correspond to the NBER-dated contractions (from peak to trough). Columns 3-6 correspond to the decomposition as described in (2).

1983 level is driven entirely by changes in the labor outcomes of different groups.

#### Out of Labor Force

Panel C of Figure 6 plots the two counterfactual series of the 2010 OLF-to-population ratio against its actual 2010 value. We highlight two features of the figure: (1) The actual 2010 OLF-to-population ratio is always higher than the counterfactual series calculated by holding the labor outcomes of different groups at their 2010 level; and (2) prior to 1994, the actual 2010 OLF-to-population ratio is always higher than the counterfactual series that is calculated by holding the sociodemographic composition constant at its 2010 level, although after 1994 the counterfactual is sometimes higher. Thus, we infer that the demographic composition contributes substantially to the high OLF-to-population ratio in 2010.

Table 6 and Figure 7, Panel C formalize this result, showing that the total change in the incidence of the OLF-to-population ratio between any year prior to 2009 and 2010 is positive. It also shows that the change in the sociodemographic composition contributes to a higher OLF-to-population ratio. The contribution from the change in the sociodemographic composition has been increasing since 1968. The contribution from the change in the labor outcomes of different groups has been significantly smaller, and has even lowered the OLF-to-population ratio in more recent years. Turning our attention once again to the 1983 and 2010 comparison, we see that the OLF-to-population ratio increased 3.68 percentage points from 1983–2010, of which 1.66 percentage points can be attributed to the change in the demographic composition. Thus, changes to the demographic composition of the population have played a large role in increasing the 2010 OLF-to-population ratio relative to its 1983 level.

#### 4. FORECAST OF THE OLF-TO-POPULATION RATIO

Our findings show that changes in the employment- and OLF-to-population ratios of 25–64-year-old men during the last four decades are, to a large degree, associated with changes in the sociodemographic composition of the population. Using the labor outcomes of different sociodemographic groups from 2010 and a projected sociodemographic composition of the population in 2015, we are able to create projections of the aggregate labor outcomes in 2015 using equation (1).

As Figure 1 shows, there is a large cyclical component in the employmentand unemployment-to-population ratios. Consequently, the forecasts of these ratios depend heavily on the business cycle phase of the year of our decomposition. In addition, the changes in unemployment (and employment to a lesser extent) are mostly dominated by changes in the labor outcomes of different

Panel A: Change Decomposition - Employment Percentage Point Change 1970 1980 1990 2000 2010 Year Panel B: Change Decomposition - Unemployment Total Change Change Because of Demographics Percentage Point Change 1970 1980 1990 2000 2010 Year Panel C: Change Decomposition - OLF Total Change Change Because of Demographics Percentage Point Change 1980 1990 2000

Figure 7 The Decomposition of the Change in Labor Outcomes

Notes: Authors' own calculations from the IPUMS-CPS data.

sociodemographic groups rather than changes in the sociodemographic composition. In contrast, the OLF-to-population ratio has a much smaller cyclical component. Consequently, we focus on forecasting the OLF-to-population ratio.

Year	Total Change		Because of Outcomes	Change Because of Sociodemographics					
		<b>Percentage Points</b>	% of Total Change	Percentage Points	% of Total Change				
1968	6.87	7.04	102.42	-0.17	-2.42				
1969	7.20	7.31	101.46	-0.10	-1.46				
1970	6.25	6.28	100.57	-0.04	-0.57				
1971	5.10	5.18	101.62	-0.08	-1.62				
1972	5.41	5.44	100.53	-0.03	-0.53				
1973	5.90	5.88	99.61	0.02	0.39				
1974	6.02	5.97	99.14	0.05	0.86				
1975	3.00	2.89	96.20	0.11	3.80				
1976	3.99	3.83	95.98	0.16	4.02				
1977	4.22	4.05	95.79	0.18	4.21				
1978	5.24	5.07	96.64	0.18	3.36				
1979	5.48	5.28	96.33	0.20	3.67				
1980	4.66	4.49	96.41	0.17	3.59				
1981	3.73	3.72	99.76	0.01	0.24				
1982	1.86	1.86	99.75	0.00	0.25				
1983	0.18	0.18	101.02	0.00	-1.02				
1984	2.82	2.86	101.41	-0.04	-1.41				
1985	3.50	3.56	101.66	-0.06	-1.66				
1986	3.40	3.55	104.50	-0.15	-4.50				
1987	3.62	3.78	104.28	-0.15	-4.28				
1988	4.32	4.53	105.07	-0.22	-5.07				
1989	4.72	4.95	104.72	-0.22	-4.72				

Table 5 (Continued) Decomposition of the Change in the Unemployment-to-Population Ratio Between 2010 and Earlier Years

Year	Total Change		Because of Outcomes	Change Because of Sociodemographics				
	J	Percentage Points	% of Total Change	Percentage Points	% of Total Change			
1990	4.69	4.96	105.69	-0.27	-5.69			
1991	2.90	3.25	112.03	-0.35	-12.03			
1992	2.26	2.58	114.43	-0.33	-14.43			
1993	2.45	2.67	108.79	-0.22	-8.79			
1994	3.55	3.79	106.79	-0.24	-6.79			
1995	4.47	4.63	103.61	-0.16	-3.61			
1996	4.28	4.49	105.03	-0.21	-5.03			
1997	4.78	5.00	104.59	-0.22	-4.59			
1998	5.32	5.49	103.31	-0.18	-3.31			
1999	5.83	5.90	101.11	-0.06	-1.11			
2000	5.99	6.00	100.09	-0.01	-0.09			
2001	5.57	5.53	99.21	0.04	0.79			
2002	4.16	4.12	98.96	0.04	1.04			
2003	3.93	3.87	98.53	0.06	1.47			
2004	4.24	4.20	98.87	0.05	1.13			
2005	4.89	4.87	99.63	0.02	0.37			
2006	5.24	5.23	99.98	0.00	0.02			
2007	5.25	5.19	98.87	0.06	1.13			
2008	4.76	4.71	99.04	0.05	0.96			
2009	0.83	0.80	96.35	0.03	3.65			
2010	0.00	0.00	0.00	0.00	0.00			

Notes: Authors' own calculations from the IPUMS-CPS data. Bold rows correspond to the NBER-dated contractions (from peak to trough). Columns 3-6 correspond to the decomposition as described in (2).

To perform this forecasting exercise, we project the sociodemographic composition of the population of 25–64-year-old men in 2015. In projecting this composition, we focus on the changes in age of the individuals in the 2010 sample while holding education, race, and marital characteristics constant at their 2010 levels.

To simulate a sample of 25–64-year-old male workers in 2015, we use the 2010 sample of 20–59-year-old male workers and construct the age variable for 2015. We use the age-specific annual male mortality rates from the Social Security Administration<sup>4</sup> and accordingly choose which workers of a particular age survive from 2010–2015.<sup>5</sup> Each worker in the simulated sample is aged five years, but has the same education, race, and marital status as in the 2010 sample. We use the projected 2015 population and the CPS sampling weights to construct the sociodemographic composition terms in equation (1). Then we use these forecasted demographic composition terms and the labor outcomes of the corresponding sociodemographic groups from 2010 to construct the predicted aggregate 2015 OLF-to-population ratio using equation (1).

Note that this exercise assumes that the mortality rates for each age remain unchanged from 2007–2015. Also, we use the sampling weights from 2010, which may not deliver a representative population for our simulated 2015 sample (for example, because we do not adjust the weights to reflect the demographic composition of the surviving individuals). However, given the relatively short forecast horizon, these weights provide a good approximation for aggregation. Finally, when aging the 2010 population, we do not accordingly adjust demographic factors other than age. For example, aging a male from 20 to 25 might alter both his educational attainment and marital status. Our forecasting exercise does not take these effects into account.

The results of our forecast are displayed in Table 7. Panel C of Table 7 contains our forecast and the U.S. Census forecast for the age distribution of 25–64-year-old men in 2015. As the forecast shows, the shares of 55–64 and 25–34-year-old males are projected to increase, while the share of 35–54-year-old males is projected to decrease.

Panel A of Table 7 displays the results of the forecast of the OLF-to-population ratio based on the labor status outcomes of different groups in 2010. For comparison, Panel B contains the results based on the labor status outcomes of different groups in 2007, i.e., the year of a recent business cycle peak. The results show that under both sets of labor status outcomes of different groups, the OLF-to-population ratio is predicted to reach more than 16 percent in 2015 as compared to the actual rate of 14.7 percent in 2010.

<sup>&</sup>lt;sup>4</sup> Data available at: www.ssa.gov/oact/STATS/table4c6.html#ss.

<sup>&</sup>lt;sup>5</sup> For example, the probability that a 20-year-old worker in 2010 survives to 2015 is  $(1 - p_{20}^m)(1 - p_{21}^m)(1 - p_{22}^m)(1 - p_{23}^m)(1 - p_{24}^m)$ , where  $p_a^m$  is the annual mortality rate of a worker at age a.

Table 6 Decomposition of the Change in the OLF-to-Population Ratio Between 2010 and Earlier Years

Year	Total Change		Because of Outcomes	Change Because of Sociodemographics				
	J	<b>Percentage Points</b>	% of Total Change	<b>Percentage Points</b>	% of Total Change			
1968	8.65	7.92	91.51	0.73	8.49			
1969	8.32	7.55	90.73	0.77	9.27			
1970	8.21	7.41	90.25	0.80	9.75			
1971	7.65	6.84	89.40	0.81	10.60			
1972	7.05	6.15	87.27	0.90	12.73			
1973	6.50	5.51	84.88	0.98	15.12			
1974	6.20	5.28	85.16	0.92	14.84			
1975	5.26	4.31	81.82	0.96	18.18			
1976	4.61	3.50	75.80	1.12	24.20			
1977	4.85	3.70	76.15	1.16	23.85			
1978	4.55	3.46	76.03	1.09	23.97			
1979	4.83	3.60	74.62	1.22	25.38			
1980	4.53	3.27	72.15	1.26	27.85			
1981	4.18	2.85	68.10	1.33	31.90			
1982	3.97	2.50	62.90	1.47	37.10			
1983	3.68	2.06	55.89	1.62	44.11			
1984	3.43	1.77	51.70	1.66	48.30			
1985	3.31	1.68	50.69	1.63	49.31			
1986	3.16	1.46	46.24	1.70	53.76			
1987	3.55	1.72	48.30	1.84	51.70			
1988	3.35	1.53	45.69	1.82	54.31			
1989	3.54	1.63	46.01	1.91	53.99			

Table 6 (Continued) Decomposition of the Change in the OLF-to-Population Ratio Between 2010 and Earlier Years

Year	Total Change	· ·	Because of Outcomes	Change Because of Sociodemographics				
	_	Percentage Points	% of Total Change	Percentage Points	% of Total Change			
1990	3.51	1.51	43.12	2.00	56.88			
1991	3.38	1.54	45.67	1.84	54.33			
1992	3.36	1.42	42.26	1.94	57.74			
1993	2.87	0.93	32.32	1.94	67.68			
1994	1.73	-0.28	-16.19	2.02	116.19			
1995	2.15	-0.03	-1.41	2.18	101.41			
1996	2.21	0.18	7.96	2.03	92.04			
1997	2.47	0.54	21.98	1.93	78.02			
1998	2.51	0.63	25.06	1.88	74.94			
1999	2.21	0.37	16.76	1.84	83.24			
2000	2.23	0.36	15.93	1.88	84.07			
2001	2.05	0.35	16.92	1.70	83.08			
2002	1.72	0.19	10.93	1.53	89.07			
2003	1.14	-0.24	-20.80	1.38	120.80			
2004	0.67	-0.53	-78.90	1.20	178.90			
2005	0.70	-0.27	-38.66	0.97	138.66			
2006	0.76	0.00	-0.16	0.76	100.16			
2007	0.95	0.27	28.02	0.69	71.98			
2008	0.52	0.08	14.69	0.44	85.31			
2009	-0.15	-0.43	284.63	0.28	-184.63			
2010	0.00	0.00	0.00	0.00	0.00			

Notes: Authors' own calculations from the IPUMS-CPS data. Bold rows correspond to the NBER-dated contractions (from peak to trough). Columns 3-6 correspond to the decomposition as described in (2).

Table 7 Predicted Aggregate OLF-to-Population Ratio Among 25–64-Year-Old Men, 2015

Danal A. B	ased on 2010 La	ahar Outcomes	
I allei A. B	Actual 2010 La	Predicted 2015	
Employment-to-Population	76.33	X	
Unemployment-to-Population	8.92	X	
OLF-to-Population	14.75	16.27	
Panel B: B	ased on 2007 La	abor Outcomes	
	Actual 2007	Predicted 2015	
Employment-to-Population	82.53	X	
Unemployment-to-Population	3.67	X	
OLF-to-Population	13.79	16.04	
Panel C	: Age Composition	on (Percent)	
Age	Actual 2010	Simulated 2015	Census 2015
25–34	25.80	26.36	26.74
35–44	25.20	23.93	24.63
45–54	27.50	25.87	25.43
55–64	21.50	23.85	23.20

#### 5. CONCLUSIONS

The OLF-to-population ratio among 25–64-year-old men has increased from 6.5 percent in 1970 to 14.7 percent in 2010. In the aftermath of the 1969–1970 recession, the employment-to-population ratio among this group was 89.1 percent, while in the aftermath of the 2007–2009 recession, the ratio is nearly 13 percentage points lower. Throughout this article we have examined the degree to which these changes can be explained by changes in the composition of the population by age, race, education, and marital status, and the degree to which they can be attributed to changes in the labor market outcomes of different sociodemographic groups.

We find that the rise in the OLF-to-population ratio since the early 1980s is primarily a result of changes in the demographic composition of the population. Changes in the demographic composition account for about 25 percent of the increase in the employment-to-population ratio during the same period, and changes in the unemployment-to-population ratio are almost entirely driven by changes in the employment status composition. Finally, simulating the 2010 sample five years forward and using labor outcomes of different sociodemographic groups from 2010, we project that the OLF-to-population ratio among 25–64-year-old men will rise to 16 percent in 2015.

#### REFERENCES

- Autor, David H., and Mark G. Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *The Quarterly Journal of Economics* 118 (February): 157–205.
- Fallick, Bruce, and Jonathan Pingle. 2006. "A Cohort-Based Model of Labor Force Participation." Federal Reserve Board of Governors *Finance and Economics Discussion Series* 2007–2009.
- Juhn, Chinhui, Kevin M. Murphy, and Robert H. Topel. 1991. "Why Has the Natural Rate of Unemployment Increased Over Time?" *Brookings Papers on Economic Activity* 22 (2): 75–142.
- Juhn, Chinhui, Kevin M. Murphy, and Robert H. Topel. 2002. "Current Unemployment, Historically Contemplated." *Brookings Papers on Economic Activity* 33 (1): 79–136.
- Little, Allan, and Steve Bradley. 2007. "The Determinants of Changes in the Risk of Unemployment and Inactivity in Britain: A Decomposition Analysis." Labour Economics Research Group Discussion Paper No. 17.
- Polivka, Anne E., and Stephen M. Miller. 1995. "The CPS After the Redesign: Refocusing the Economic Lens." Manuscript, Bureau of Labor Statistics.
- Shimer, Robert. 1999. "Why is the U.S. Unemployment Rate so Much Lower?" In *NBER Macroeconomics Annual 1998*, Volume 13, edited by Ben S. Bernanke and Julio Rotemburg. Chicago, The University of Chicago Press, 11–61.
- U.S. Census Bureau. 2008. "2008 National Population Projections." Washington, D.C.: U.S. Census Bureau. www.census.gov/population/www/projections/2008projections.html.

## Strategic Behavior in the Tri-Party Repo Market

Huberto M. Ennis

Repo contracts are a kind of collateralized loan that has become predominant in the United States among large cash investors. There are several types of repo contracts, such as bilateral delivery-versus-payment repos, interdealer repos, and tri-party repos. A significant portion of repo transactions in the United States take the form of tri-party repos, where a third party (a clearing bank) provides collateral management and settlement services to the borrower and the lender. The tri-party segment of the U.S. repo market is the subject of this article.

The tri-party repo market played a significant role during the 2007–2009 global financial crisis. Tri-party repos were, for example, a major source of secured funding for Bear Sterns prior to its demise. In March 2008, repo lenders in general, and tri-party repo counterparties in particular, lost confidence in their ability to recoup loans to Bear Stearns and, hence, refused to renew them, asking instead for immediate repayment (Bernanke 2008). To avoid a failure, the Federal Reserve facilitated the acquisition of Bear Stearns by the bank J.P. Morgan Chase. The withdrawal of tri-party repo funding also played a role in the collapse of Lehman Brothers in September 2008. As a result of the events during the crisis, it is now widely believed that the tri-party repo market is subject to serious vulnerability (see, for example, Dudley [2009]). Attesting to this is the fact that in 2009 the New York Fed asked a group of senior private U.S. bank officials to form a task force "to address the weaknesses" in the

I would like to thank Jeff Lacker for many stimulating conversations that motivated me to write this article, Jim Peck for answering my game theory questions patiently, and Borys Grochulski, Bob Hetzel, Andreas Hornstein, Tim Hursey, Todd Keister, Antoine Martin, Ned Prescott, Alex Wolman, and John Walter for comments on an earlier draft. All errors and imprecisions are of course my exclusive responsibility. The views expressed in this article are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: huberto.ennis@rich.frb.org.

infrastructure of the tri-party repo market (Federal Reserve Bank of New York 2010). A broad set of reforms are currently under way.<sup>1</sup>

In this article, we study a simple model of the tri-party repo arrangement that allows us to capture in a parsimonious way some of the strategic interactions that arise in this market. In our analysis, we use standard non-cooperative game theory to uncover the basic mechanisms that can create some of the vulnerabilities commonly attributed to the tri-party repo market. We will show that a change in perceptions can create a sudden coordinated withdrawal of lenders from this market. Also, we will highlight the crucial role that the clearing bank plays in this game of "withdrawing before the rest," which appears to be a good representation of the situation that was present in the tri-party repo market during the recent financial crisis.

A repo (repurchase agreement) transaction is a sale of an asset that is combined with an agreement to repurchase the asset at a pre-specified price on a later day. Effectively, it is equivalent to a collateralized loan, where the loan is the amount paid for the initial sale and the asset plays the role of collateral. Repayment of the loan takes place at the repurchase time, with the interest rate being implicit in the repurchase price. In a tri-party repo, a third party—the tri-party agent—facilitates the transaction between the two main parties, the lender (a cash investor such as a money market mutual fund) and the borrower (a securities dealer such as the broker-dealer arm of an investment bank). The tri-party agent provides custodial and other services to the lender and efficient collateral assignment and allocation tools to the borrower. Settlement happens entirely in the books of the tri-party agent where both the borrower and the lender have cash and securities accounts. Also, in many cases, the tri-party agent (via the so-called "morning unwind") extends intraday credit to the borrowers to give them access, during the day, to the securities used as collateral overnight. In the United States, the tri-party agents are the two clearing banks, Bank of New York Mellon and J.P. Morgan Chase.2

The volume of repo transactions in the United States is large. There are no official data covering the entire market but Gorton and Metrick (forthcoming) estimate that its size peaked before the crisis at a level that is in the same order of magnitude as the value of all the assets held by U.S. commercial banks (approximately \$12 trillion). The tri-party repo segment of the market is large as well. The value of securities financed in this way was around \$1.7 trillion at the end of 2011, down from about \$2.8 trillion in early 2008 (Federal

<sup>&</sup>lt;sup>1</sup> Implementation of the reforms have proven to be more difficult than previously expected. On February 15, 2012, the New York Fed issue a statement indicating that the vulnerabilities in this market still persist.

<sup>&</sup>lt;sup>2</sup> In what follows, for the purpose of concreteness, we will always call the lender in the triparty repo the (cash) investor, and we will call the borrower the (securities) dealer. The tri-party agent will be called the clearing bank, or sometimes just the bank, for short.

Reserve Bank of New York 2010). Furthermore, some large broker dealers, arguably of systemic importance, finance large portions of their portfolios in this market. While U.S. Treasury securities and agency mortgage-backed securities (considered virtually riskless) are the most common class of assets used as collateral in tri-party repos, equities and other fixed income securities are also sometimes used. According to some estimates, at its peak in early 2008, about 30 percent of the assets used as collateral were subject to non-negligible liquidity risk (Federal Reserve Bank of New York 2010).<sup>3</sup>

The amount of the loan in a repo transaction is often lower than the value of the posted collateral. In other words, the value of the collateral gets discounted and this discount is commonly referred to as a "haircut." Haircuts are aimed at reducing the exposure of the lending side to liquidation costs in case the borrower defaults (see Gorton and Metrick 2010). In principle, choosing the appropriate haircut would leave the repo transaction free of virtually any repayment risk. This is the case because repo transactions are generally exempt from the automatic stay that applies to debt under the U.S. Bankruptcy Code. This implies that the lender side in a repo transaction can take possession of the collateral immediately upon failure of the borrower.<sup>4</sup>

Data on haircuts for different types of repos is limited. However, the available evidence suggests that the level and sensitivity of haircuts depend on the kind of repo transaction being considered. Gorton and Metrick (2010, forthcoming), for example, study a sample of interdealer repo transactions and show that the average haircut increased significantly during the crisis. This is the manifestation of what they call "the run on repos." Repos were used to finance portfolios of securities and, as the haircuts increased, the capacity to borrow against those securities decreased. The owners of the securities, then, had to find alternative sources of funding or sell the securities in the market. This deleveraging is tantamount to the liquidation of loans that takes place in traditional bank runs.

In contrast, Copeland, Martin, and Walker (2010) show that collateral haircuts in the tri-party repo market did not appear to adjust in any meaningful way to changes in the riskiness of the borrowers. The infrastructure that made tri-party repos attractive to investors seems to have made it less convenient for them to adjust collateral haircuts on a per-transaction basis. Instead, when the financial conditions of a given dealer deteriorated, cash investors tended to withdraw from dealing with such a dealer (PRC Task Force 2010). Evidence

<sup>&</sup>lt;sup>3</sup> Up-to-date information on the composition of collateral in the tri-party repo market can be found at www.newyorkfed.org/tripartyrepo/.

<sup>&</sup>lt;sup>4</sup> The repo exemption from the stay is likely to extend to the case of the failure of a broker dealer, as explained by Copeland, Martin, and Walker (2010, Appendix C). There is an ongoing debate about the appropriateness of granting safe-harbor exemptions from the automatic stay to a broad range of derivative transactions, including repos. See, for example, Roe (2009) and Lubben (2010).

suggests that this behavior was predominant during the events that led to the failure of Lehman Brothers (Copeland, Martin, and Walker 2010). This way of reacting to counterparty credit risk in the tri-party repo market is taken as a premise in this article and plays an important role in the theoretical arguments advanced later. In particular, we will investigate the problems that can arise in the strategic interaction between the main players in this market given that withdrawal from lending (and not adjustments of haircuts) constitutes the typical reaction to a change in perceptions about the viability of the borrowing side in the transaction.

Policymakers believe that a breakdown of the repo market can have systemic consequences. In March 2008, after the collapse of Bear Stearns, the Federal Reserve created the Primary Dealer Credit Facility (PDCF) on the premise that "unusual and exigent" circumstances justified the provision of emergency (collateralized) lending to large securities dealers. The idea behind the PDCF was to provide backup liquidity to dealers to give them time to arrange other sources of funding if repo lenders were to suddenly withdraw from the market. The program was designed as a backstop facility, charging a penalty rate on tri-party repo transactions in which the Fed took the lending side (see Adrian, Burke, and McAndrews 2009). Initially, only high-quality collateral (investment-grade securities) was accepted in the PDCF. At the time of the failure of Lehman Brothers, the Fed expanded collateral acceptability to a broader set of assets and usage of the PDCF soared. We will use our model to illustrate one possible role for a lending facility such as the PDCF. However, a more careful assessment of the suitable policy responses to the type of vulnerabilities highlighted in this article is left for future research.

The article is organized as follows. In the rest of this section, we describe the "morning unwind," a feature of the tri-party repo market that is crucial for understanding the main strategic interaction explored in this article. In Section 1, we set up a simple model of the tri-party repo market and proceed to study the induced strategic interaction between investors and the clearing bank using standard tools in non-cooperative game theory. In Section 2, we discuss some related issues that pertain to the functioning of the tri-party repo market as presented in this article. Finally, Section 3 concludes.

#### The Morning Unwind

The maturity of most tri-party repo contracts is overnight, but there are also contracts being arranged for a week, 30 days, and even longer periods of time. A common practice in this market, however, is that the clearing bank "unwinds" all repos, regardless of maturity, at the beginning of each day (at around 8:00 a.m. EST).

The process of unwinding takes place as follows. Overnight, the cash investor has the securities in its account at the clearing bank. As part of an implicit arrangement, early in the morning (before the open of Fedwire

securities at 8:30 a.m. EST), the clearing bank transfers the securities back from the investor's account to the dealer's account, and transfers the corresponding cash to the investor (much like in a cancellation of the repo). To finance the transfer of cash, the clearing bank (normally) extends intraday credit to the dealer. In other words, the investor gets a credit in its cash account at the bank and the dealer gets a debit, which usually results in an intraday overdraft of its cash account.

There are several reasons why it is convenient for investors and dealers to have the repos unwound in the morning. Investors benefit from having their cash available to make various payments and to satisfy withdrawal demands placed by their clients during the day. Dealers benefit from having access to the securities for the purpose of trading. In fact, as a result of the trading activities of dealers, the composition of their portfolio of securities changes during the day. If some of the securities being used as collateral in outstanding repos are sold, then they need to be substituted with new securities. This process of *collateral substitution* is simpler if all the securities are transferred to the dealer's account in the morning and only reallocated back to repo contracts at the end of the day.

With the morning unwind, the tri-party repo contract constitutes a loan based on the combination of two sources of funding: investors covering the night and the clearing bank covering the day. As with the overnight credit provided by investors, the intraday credit provided by the clearing bank is secured by the securities held by the dealer in its account at the bank.<sup>5</sup> In other words, if the dealer were to fail during the day, after the unwind has occurred, then the clearing bank would get ownership of the securities as a way to cancel the dealer's overdraft. If, instead, the failure of the dealer were to occur during the night, then investors would retain ownership of the securities that served as collateral for the tri-party repo transaction.

The morning unwind, then, to the extent that it is financed with the provision of intraday credit to dealers, exposes the clearing bank to the risk of receiving ownership of a batch of securities upon the failure of one (or more) of those dealers.<sup>6</sup> This unplanned increase in assets of the clearing bank may create some extra costs associated with balance sheet capacity (capital

<sup>&</sup>lt;sup>5</sup> The clearing bank has a lien on the dealer's collateral structured as a repo with broad flexibility for collateral substitution. When the dealer sells (delivery versus payment) a security during the day, the cash received as payment cancels out the part of the overdraft that is no longer collateralized because of the sale of the security. When a dealer delivers a security free of payment, the clearing bank is protected by its "right of offset" on all the securities that the dealer has at the clearing bank, including those that were not used in tri-party repo transactions.

<sup>&</sup>lt;sup>6</sup> The ongoing reorganization of the market intends to reduce the predominance of the automatic "morning unwind" practice. See PRC Task Force (2010) for details. However, in the statement issued on February 15, 2012, the New York Fed said: "the amount of intraday credit provided by clearing banks has not yet been meaningfully reduced, and therefore, the systemic risk associated with this market remains unchanged."

constraints, for example). Furthermore, it is possible that part of the overdraft extended to the dealer by the clearing bank is, in turn, being financed by an intraday overdraft of the clearing bank on its account at the Fed. If the dealer fails and the clearing bank cannot resell the securities by the end of the day, it may incur an overnight overdraft at the Fed, which is much more expensive, or it may need to borrow at the discount window. Aside from being provided at a penalty rate, discount window borrowing may also be associated with a stigma effect that can make such an activity very costly for the clearing bank.<sup>7</sup> The risk of incurring these costs is likely to be a crucial determinant of the willingness of the clearing bank to unwind the repos every morning. The clearing bank retains the right to refuse to unwind the repos of any particular dealer.

At the end of the day, tri-party repos are "rewound" and cash investors are the party exposed to the risk of failure of the dealer during the night. It is common for cash investors in tri-party repos to accept certain securities that they are not allowed to hold permanently in their portfolios. If the dealer were to fail during the night, then, the cash investor would receive a batch of securities that they would need to sell as soon as possible. Rush sales may result in unfavorable prices (beyond the haircut applied to the collateral), effectively exposing cash investors to financial losses.

It is important to realize here that the reason why the clearing bank is (potentially) exposed to credit risk during the day is not because of the process of unwinding the repos in the morning itself, but because such unwinding is generally financed with intraday credit (an overdraft) extended by the clearing bank to the dealer. If, every morning, the dealer were to have enough cash in its account at the clearing bank, then the unwinding would make the repo essentially a secured debt contract with a half-day maturity. The only exposure in that case would be on the lending side (cash investors) and only to the extent that the haircut on the collateral is not enough to cover any discount associated with selling the assets.

#### 1. A SIMPLE MODEL

The tri-party repo market in the United States is a complex system. There are multiple participants facing diverse situations. Some of them are always there, day after day, and some only participate occasionally. The clearing banks, the main broker dealers, and some of the large cash investors participate every day; one can suspect, then, that implicit relationships and reputation, for example, play a significant role in determining outcomes (Copeland, Martin, and Walker 2010). Dealing with all these different dimensions formally is a challenging

<sup>&</sup>lt;sup>7</sup> For recent work on the possibility of stigma at the discount window see Ennis and Weinberg (2010) and Armantier et al. (2011).

task and it may not be the most illuminating approach. Instead, here, we will provide a very simple environment that captures only some of the forces at play in this market and we will use standard non-cooperative game theory to analyze the strategic component associated with such a situation.<sup>8</sup>

The model is very simple. There are two time periods, t=1,2, and three types of agents, a clearing bank, a securities dealer, and N cash investors. At the beginning of period 1, each cash investor has an endowment of c dollars and the dealer has the opportunity to invest 1 dollar in securities, which will pay  $1+\rho$  at the end of period 2. We allow for  $\rho$  to be a random variable and consider the natural case in which  $\rho$  has a positive expected value. We also assume that Nc>1.

At the beginning of period 1, cash investors deposit (some of) their cash at the clearing bank. Also at that time, the dealer can request a 1 dollar intraday overdraft at the clearing bank to buy the securities. The clearing bank may or may not agree to grant the dealer's overdraft request.

At the end of period 1, the dealer needs to close the overdraft in its account at the clearing bank. We assume that overnight overdrafts are expensive enough to give the dealer incentives to do this. In order to obtain the cash needed to fund the overdraft position, the dealer arranges tri-party repos with cash investors using the securities as collateral. The interest rate on the repos is taken parametrically and denoted by r.

If the dealer is not able to repo the securities, then it has to sell the securities to pay back as much of the overdraft as possible. We assume that securities sold before the end of period 2 only return a portion of what was invested. In such a situation, then, the dealer gets no return and the clearing bank experiences a loss equal to  $y_B > 0$ .

If the dealer is able to repo the securities, it closes the overdraft at the bank, and the next morning the bank has to decide whether or not to unwind the repos. If the bank decides not to unwind the repos, then the dealer has no funding for the securities, it fails, and investors take possession of the collateral. We also assume that investors cannot hold the securities and need to sell them at a loss at the beginning of period 2. In such a case, again, the dealer gets no return and investors experience a loss equal to  $y_I > 0$ . The dealer stops being a customer of the bank at that point and the bank gets no payoff from the transaction.

If the bank agrees to unwind the repos instead, the dealer gets a new daylight overdraft in its bank account and investors get their cash and interest

<sup>&</sup>lt;sup>8</sup> See Duffie (2010) for a detailed description of the various activities generally undertaken by broker dealers and the role that the repo market plays in funding those activities.

<sup>&</sup>lt;sup>9</sup> In the United States, most tri-party repos are arranged in the morning and settle in the books of the clearing bank late in the afternoon, after the close of Fedwire securities. For the formal representation of the problem, the only relevant aspect is that, each day, new repo funding is arranged only after the morning unwind.

back. At the end of the day, the securities pay off and the revenue is used by the dealer to close the overdraft and pay a fee  $\phi$  to the bank.

Note that the initial overdraft could be thought of as the result of the unwinding of a (set of) pre-existing repo contract(s). In that sense, we could think that our simple framework is able to handle two rounds of unwinding, to the extent that the decision to unwind, in this model, will be exclusively driven by forward-looking considerations. This interpretation of the initial overdraft will be useful when we discuss some of the results.

Since we are assuming that Nc > 1, investors' initial endowment would be enough to (fully) fund the investment opportunity of the dealer. The way this funding is channeled from investors to dealers is via the clearing bank. The clearing bank receives an initial deposit  $d \le Nc$  from investors and then grants a daylight overdraft to the dealer. If d > 1, then, on the books of the clearing bank, the overdraft (loan) to the dealer is (fully) funded by the deposit of investors. However, if investors do not deposit all of their endowment at the bank and d < 1, then initial funding for the dealer could still be available. At the beginning of period 1, the bank obtains daylight credit from the central bank in the amount 1-d. Later in the period, when (and if) the dealer secures repo funding from investors, the corresponding cash that closes the negative position of the dealer can be used by the bank to close its negative position with the central bank. In this way, the bank can avoid a more expensive overnight overdraft at the central bank.

Finally, notice that we have simplified the dealer's side of the problem by assuming that whenever funding is not forthcoming, the dealer fails. This strategy allows us to concentrate our attention on the interaction between investors and the clearing bank. <sup>10</sup> Furthermore, when the dealer fails and the securities need to be liquidated before the end of period 2, the proceeds from the sale are not enough to cover the total value of the loan—the lender suffers losses. In effect, this is a direct counterpart of postulating that insufficient haircuts are applied to the collateral. As discussed in the introduction, the evidence described in Copeland, Martin, and Walker (2010) suggests that this is a reasonable approach to take.

#### The Non-Cooperative Game

The key strategic interaction in the model is between the clearing bank and the set of investors. To study the outcome from this interaction we can use the tools of non-cooperative game theory. In particular, we will concentrate our attention here on the implied formal game played between the bank and investors.

<sup>10</sup> See Martin, Skeie, and von Thadden (2010) for a more fleshed out formal treatment of the role of investors' decisions in determining the fate of the dealer.

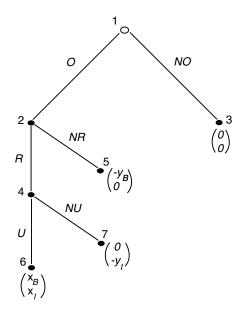


Figure 1 The Game with No Uncertainty: Game 1

Let us start with the case when N=1 and  $\rho=H\in\mathbb{R}_+$  (i.e.,  $\rho$  is a given number greater than zero, not a random variable). Assume that  $H > \phi + r$ . The extensive form representation of this game, which we call Game 1, appears in Figure 1. The game starts in node 1 (represented by an open circle in the figure) with the move by the clearing bank, who has to decide whether to grant the dealer a daylight overdraft (O) or not (NO). After that, if an overdraft is granted, the investor has to decide whether to enter a repo contract with the dealer (R) or not (NR). This is the decision presented in node 2. Finally, if a repo contract is arranged, then the bank has to decide, in node 4, whether to unwind the repo (U) or not (NU) the next morning. In each of the terminal nodes (nodes 3, 5, 6, and 7) the payoffs of the players are listed in a column, with the top element representing the payoff for the clearing bank (the first player to move) and the bottom element representing the payoff of the investor. We use the variables  $x_i$  with i = B, I to represent the payoffs to the bank (B)and the investor (I) in the case where an unwinding of the repo happens on the morning of date 2. From our description of the model above, we know that  $x_B = \phi$  and  $x_I = r$ . In a less stylized setup,  $x_i$  could be equal to something more complicated, but the basic results from the strategic interaction will be the same as long as the conditions on  $x_i$  and  $y_i$  established below still hold.

We look for a subgame perfect Nash (SPN) equilibrium of this game. Since Game 1 is a finite game of perfect information, an equilibrium always

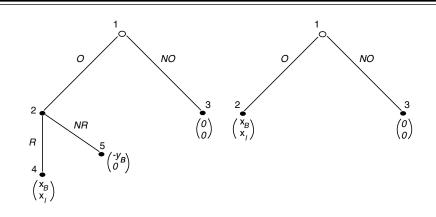


Figure 2 Solving for the SPN Equilibrium of Game 1

exists, and, given the payoffs, it is easy to see that the equilibrium is actually unique (see, for example, Osborne and Rubinstein [1994]).

**Proposition 1** There is a unique SPN equilibrium of Game 1 for which the equilibrium actions are (O, R, U).

**Proof.** As is standard with dynamic games, we proceed by solving backward. First, consider the decision of the bank in the subgame that starts at node 4, that is, after investors have agreed to repo the securities. If the bank unwinds the repos, then it gets a payoff equal to  $x_B$ , which is greater than the payoff of zero obtained from not unwinding. Then, the bank will agree to unwind the repos. We can now write an auxiliary game tree that takes this result into account. This is the tree represented in the left-hand side of Figure 2. Following the same logic, we can now solve backward in this game to find that the investor will agree to repo the securities because  $x_B > -y_B$ .

Finally, we can draw an auxiliary tree that incorporates this last result (on the right-hand side of Figure 2) and find that the bank will agree to grant an overdraft since  $x_B > 0$ . Hence, we have that the bank will always play O, then the investor will always play R, and lastly the bank will always play U, which completes the proof of the proposition.

When there is no uncertainty with respect to the long-term solvency of the dealer and there is only one cash investor (or a well-coordinated group of them), the dealer always receives funding from the clearing bank (via daylight overdraft) and from the investor (via repo transactions). There is no instability associated with the tri-party repo contract in this case.

# **Uncertainty over the Dealer's Solvency**

Suppose now that  $\rho$  is, in fact, a random variable that can take value H>1 with probability  $\xi$  and -L with probability  $1-\xi$ . We associate the outcome  $\rho=-L$  with a situation where the dealer experiences a solvency problem not triggered by the actions of the participants in the tri-party repo market. We will consider two cases: one where the game is played without the investor or the bank knowing the realization of the random variable  $\rho$ , and the other where the bank gets to know the realization of  $\rho$  before deciding whether or not to unwind the repos the morning of date 2.

#### Uninformed clearing bank

In this first case, both the bank and the investor, when making decisions, share the same degree of uncertainty about the expected performance of the dealer. The structure of the game is almost exactly the same as in Game 1, except that the payoff to the bank in terminal node 6 is now given by  $\xi x_B + (1 - \xi) (-f_B)$ , where  $f_B$  is the loss to the exposed bank when the dealer fails. We call this Game 2a. Note that the payoff to the repo investor in node 6 is still equal to  $x_I$  since the unwinding of the repos occurs as in normal circumstances in that branch of the tree. Basically, the idea is that with some probability, the bank finds out that the dealer is insolvent after unwinding the repos and hence is left with a loss equal to  $f_B = L + r$  in our model. 12

**Proposition 2** Define  $\overline{\xi_a} \equiv f_B/(x_B + f_B)$ . If  $\xi > \overline{\xi_a}$ , then there is a unique SPN equilibrium of Game 2a for which the equilibrium actions are (O, R, U).

The proof of the proposition follows the same logic as the proof of Proposition 1, so we do not repeat it here. If the probability of the dealer not experiencing a solvency problem is high enough (i.e., if  $\xi$  is high enough), then the dealer will get funding from the bank and from the cash investor. However, if the probability  $\xi$  is below the threshold value  $\overline{\xi_a}$ , then the unique SPN equilibrium has the bank playing NO in node 1 and the dealer does not obtain funding in such a situation. We could summarize this result as saying that those dealers who are perceived as "fragile" will not get funded.

 $<sup>^{11}</sup>$  See Duffie (2010) for a thorough description of the various factors that can contribute to the failure of a dealer bank.

 $<sup>^{12}</sup>$  The bank, at the time of unwinding the repos, grants an overdraft to the dealer of size 1+r. After the dealer fails, the securities pay 1-L and the bank gets the proceeds. Hence, the net loss for the bank is equal to L+r.

<sup>&</sup>lt;sup>13</sup> Recall that in game theory, an equilibrium is a property of a profile of strategies. A strategy is a complete contingent plan of play for all possible circumstances in the game, not just the ones that occur in equilibrium. For example, when  $\xi < \overline{\xi_a}$ , the equilibrium strategy of the bank is  $\{NO, NU \text{ if the investor plays } R\}$  and the equilibrium strategy of the investor is  $\{NR \text{ if the bank plays } O\}$ . In this article, we sometimes loosely describe equilibrium play by the actions taken in equilibrium, just to keep the presentation simple.

It is interesting to note that the bank plays NO when  $\xi < \overline{\xi_a}$  because it anticipates that the investor will not be willing to enter into a repo agreement at the end of the day to finance the dealer. The investor, in turn, does not agree to participate in the repo because it anticipates that the bank will not be willing to unwind the repos the next morning if the repos were outstanding. This anticipation game makes the tri-party repo market very sensitive to changes in perceptions, not just about actual weaknesses of the dealer being funded, but also about the perceptions of other players about those weaknesses.

If we interpret the initial overdraft as (possibly) the result of an unwinding of previously arranged repos, then the model says that if the clearing bank places a high probability on the eventual failure of the dealer *the next day*, the refusal to unwind will take place immediately. This result suggests that the situation can potentially unravel long before the actual failure of the dealer is expected to occur, even if such failure is only regarded as a possibility (and not a certainty).

A crucial issue left unexplored here is how the perception of the probability of failure gets determined and how it changes over time. What the theory here makes clear is that, once such probability has crossed a certain threshold, the whole tri-party repo arrangement is bound to immediately collapse.

#### Informed clearing bank

The second case we would like to consider in this section is the case when the bank gets to know the realization of  $\rho$  before deciding whether or not to unwind the repos on the morning of date 2. We refer to this game as Game 2b. The extensive form representation of this game is provided in Figure 3 where nature moves at node 4. We denote by NF the situation when the realized state of nature is such that  $\rho = H$ , and by F the situation when  $\rho = -L$ . The other new piece of notation in Figure 3 is the payoff  $f_I$ , which is the loss experienced by the repo investor when the repo is not unwound by the bank and  $\rho = -L$ . In principle,  $f_I$  could be different than  $y_I$  because the liquidation value of the securities may depend on the state of nature.

**Proposition 3** Define  $\overline{\xi_b} \equiv f_I/(x_I + f_I)$ . If  $\xi > \overline{\xi_b}$ , then there is a unique SPN equilibrium of Game 2b for which the equilibrium actions are (O, R, U) if  $\rho = R$ , NU if  $\rho = -L$ .

**Proof.** First note that in the proper subgame that starts at node 6, the bank should agree to unwind the repos, and in the one that starts in node 7, the

<sup>&</sup>lt;sup>14</sup> Copeland, Martin, and Walker (2010) call this strategic interaction "the hand-off of risk between investors and clearing banks."

<sup>&</sup>lt;sup>15</sup> Osborne and Rubinstein (1994, 101) call games with this structure extensive games with perfect information and chance moves.

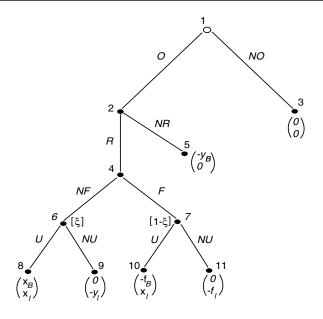


Figure 3 The Game with Uncertainty: Game 2b

bank should not unwind the repos. Now, using backward induction, we can construct the reduced game where nodes 6 and 7 are terminal nodes and the payoffs are the ones associated with nodes 8 and 11 in the full game. Given that nature moves according to the probability  $\xi$ , we have that the payoff for the investor from playing R is equal to  $\xi x_I + (1 - \xi)(-f_I)$ . Also, the payoff for the bank after playing O and given that the investor is playing R is  $\xi x_B$ . Now, again, using backward induction, we can construct a reduced game with node 4 as a terminal node and the associated payoffs being  $\{\xi x_B, \xi x_I + (1 - \xi)(-f_I)\}$ . Clearly, if  $\xi > \overline{\xi_b}$ , the investor wants to play R and, given this, the bank wants to play O (since  $\xi x_B > 0$ ).

If  $\xi < \overline{\xi_b}$ , the investor will want to play *NR* when node 2 is reached and, anticipating this, the bank will want to play *NO*. Thus, if  $\xi < \overline{\xi_b}$ , the dealer will not obtain the initial overdraft funding from the bank and no repo will be ultimately arranged.

The logic behind these results is clear. The cash investor anticipates that the bank will be able to infer somehow, before the unwinding of the repos, the future performance of the dealer. If the investor believes that it is very likely that the bank will find out that the dealer is bound to fail (and hence that the bank will not unwind the repos), then the investor will not be willing to agree to the repo transaction. In turn, anticipating this, the bank will not

grant an initial overdraft to the dealer and the whole tri-party repo arrangement collapses.

Here, again, we can loosely interpret the initial overdraft as the result of unwinding previously arranged tri-party repos. In this informal interpretation, the crucial element for such a story to work is that there must have been a change in perceptions about the situation of the dealer after repo contracts were arranged prior to the beginning of Game 2b. In particular, right at the beginning of Game 2b, it must be the case that all the participants in the triparty repo arrangement realized that the dealer actually has a probability of success (the next day) smaller than the threshold  $\overline{\xi_b}$  and that the bank will be able to find out whether or not the dealer will fail before the unwinding takes place the following day. If this is the case, then the tri-party repo arrangement *immediately* collapses, not at the time when the failure of the dealer is expected to occur but when the perceptions about that failure actually change (which could very well be much sooner, as the game illustrates).

#### Discussion

It is interesting to compare the results in Propositions 2 and 3. Note that the thresholds are increasing in the size of the loss if the dealer fails, and they are decreasing in the size of the gain if funding is granted and the dealer does not fail. This is true for both thresholds, although in Proposition 2 the relevant payoffs are those of the bank and in Proposition 3, those of the cash investor. The reason for this difference is the fact that in Game 2a the bank is playing the role of creditor at the time when the dealer fails, while in the case of Game 2b the bank finds out whether or not the dealer will fail before unwinding the repos, and if the dealer is actually expected to fail, then investors will be the party exposed to losses.

This difference in the threshold values has implications for the relationship between fragility and information flows in the tri-party repo market. We can interpret a situation with a lower threshold value as a situation where the tri-party repo arrangement is more likely to survive shifts in participants' perceptions. The idea is that the creditor will accept to stay in the transaction even after larger increases in the perceived probability of failure  $1 - \xi$  when the threshold value is lower. Then, if we think that cash investors have less to gain from the repo contract and more to lose relative to the bank—so that the threshold  $\overline{\xi}_b > \overline{\xi}_a$ —a situation where everybody anticipates that the bank will be able to obtain information about the solvency conditions of the dealer before the morning unwind (as in Proposition 3) would result in a more fragile tri-party repo market. In such a situation, it is worth noticing, increasing the haircuts applied to the collateral will tend to reduce the loss  $f_I$ , reduce the threshold value  $\overline{\xi}_b$ , and, in this way, improve the stability of the repo market.

In the simple formal game we have studied in this section, the initial perceptions about the probability  $\xi$  are shared by all participants and are correct

in the sense of being equal to the actual objective probability associated with the random variable  $\rho$ . This stark information structure hides the fact that the crucial driver of behavior in this strategic situation is the perception that the bank has about the perception of investors about the probability of failure of the dealer. Notice that, in fact, the bank would be willing to grant the initial overdraft to the dealer regardless of the bank's perception of the probability  $\xi$ , as long as the bank expects that investors will be willing to repo the securities later in the day. Whether or not investors will be willing to repo the securities depends only on the perception that those investors (and not the bank) have about  $\xi$ . So, if the bank thinks that investors are optimistic about the dealer, then, even if the bank is not, the bank will be willing to grant the initial overdraft. This is the case because the bank will get to know whether or not the dealer will fail before unwinding the repos in the morning of the second date and, hence, can effectively get out of the deal without experiencing any losses.

We have considered here the case of only one cash investor with no interim information. However, it would be more realistic to have many investors, each getting some partial information about the solvency condition of the dealer. Because the clearing bank observes the actions of investors in the tri-party repo market, it has a vantage point to aggregate all the dispersed information available to investors and hence, to some degree, anticipate the potential failure of the particular dealer. In other words, after the round of repos during the day, the bank is likely to become better informed about the situation of the dealer. The structure of Game 2b attempts to capture the gist of this situation by having the bank become *perfectly* informed before deciding whether or not to undertake the morning unwind.

Having more than one investor makes the game more complicated and can produce other interesting insights. In particular, the issue of coordination among multiple investors is key to understanding the sources of possible fragility in the tri-party repo market. We discuss some of those issues in the following sections. The analysis in this section applies to a situation where investors can (somehow) perfectly coordinate their actions and play *R* whenever such a move benefits all of them.

Before we move on to discuss potential coordination issues, it is worth mentioning an interesting implication coming out of the structure of Game 2b. In situations such as the one captured by the timing in that game, any measure aimed at reducing the potential losses of a clearing bank will not change the resiliency of the tri-party repo market. If the clearing bank (by obtaining independent information or by inferring information from the behavior of investors) can (fully) anticipate the failure of any particular dealer before the morning unwind, then the bank is effectively not exposed to actual losses (i.e., the value of  $f_B$  is irrelevant for equilibrium, as long as it is positive).

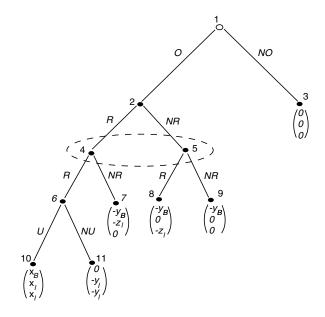


Figure 4 Coordination in the Repo Market: Game 3

Hence, any attempt at reducing a clearing bank's potential losses will not have a material effect on the behavior of the market.

# Coordination in the Repo Market

Suppose that there are N=2 cash investors and that, at the beginning of date 2, these investors play a simultaneous move game to decide whether or not to agree to enter repo contracts with the dealer. Also assume that if only one of the two investors agrees to a repo, then the dealer stops operations and the investor that entered the repo agreement experiences a loss equal to  $z_I$ . The extensive form representation of this game, which we call Game 3, is given in Figure  $4.^{16}$ 

The encircled decision nodes 4 and 5 constitute a single information set for the investor moving in those nodes. This is the result of the fact that investors play simultaneously and, hence, each investor does not know if the other investor has played *R* or *NR* at the time that he has to decide what to play (that is, the investor does not know if he is in node 4 or in node 5, respectively). As before, we look for a SPN equilibrium of Game 3.

<sup>&</sup>lt;sup>16</sup> Osborne and Rubinstein (1994, 102) call games with this structure extensive games with perfect information and simultaneous moves.

Figure 5 Normal Form Representation of the Coordination Game in the Repo Market

	Investor 2		
	R	NR	
R Investor 1	$(x_1,x_1)$	(-z <sub>1</sub> ,0)	
NR	(0,-z <sub>1</sub> )	(0,0)	

**Proposition 4** There are two pure-strategy SPN equilibria of Game 3; in one the dealer gets funded and in the other it does not.

**Proof.** Note that the branch of the game tree that starts at node 6 is indeed a proper subgame of this game. Clearly, if play reaches node 6, then the bank should agree to unwind the repos (i.e., play U) at that point. Using backward induction, we can substitute the payoff from node 10 at node 6 and consider the reduced game that results after this first iteration. In this reduced game (and in the complete game), there is one proper subgame that starts at node 2. In fact, this subgame has the structure of a coordination game between investors and has two pure-strategy Nash equilibria: (R, R) and (NR, NR) (Figure 5 depicts the normal-form representation of this coordination game).

As a result of this multiplicity, the full game actually has two pure-strategy SPN equilibria, one where investors play (R, R) if the proper subgame starting at node 2 is reached, and another where investors play (NR, NR) if this subgame is reached. In the first case, when both investors agree to enter repo transactions, the bank will be willing to grant an overdraft (i.e., play O) in node 1. The equilibrium actions will then be  $(O, \{R, R\}, U)$  and the equilibrium payoffs will be  $(x_B, x_I, x_I)$ .

In the other case, when investors play (NR,NR), we have that the bank will not agree to initially grant the overdraft and the equilibrium payoffs are equal to zero for all players since the dealer does not get funded from the outset.

The equilibrium in which the bank does not agree to grant the dealer an overdraft in node 1 captures in a stylized way a source of potential fragility in the tri-party repo market. If the clearing bank expects that, because of what amounts to a coordination failure, cash investors in the afternoon will not be willing to fund the securities dealer via repo transactions, then the bank will not be willing to grant an overdraft to the dealer in the morning. Recall that, for all practical purposes, the overdraft could originate on an initial request

for funding by a dealer or as the result of the unwinding of outstanding repo transactions. In this sense, then, the model underscores the fragility associated with the daily unwinding of repo transactions that are financed with daylight overdrafts on the accounts that securities dealers have at their clearing banks.

Note here that all agents in the model prefer that the equilibrium in which the dealer gets funded be played at all times. However, because of the possibility of a coordination failure among investors, it is consistent with rational play and equilibrium that the dealer not be funded. Martin, Skeie, and von Thadden (2010) call such a situation a repo run. One way to deal with this problem would be to have the central bank provide backstop liquidity in the repo market, as the Federal Reserve did with the PDCF. In such a situation, investors would get payoff  $x_I$  from choosing R, independent of what the other investor is choosing. This change in the structure of payoffs makes (R, R) the unique equilibrium of the game, and the dealer always gets funded. The key to this result is that the policy intervention changes the game among investors so that it is no longer a coordination game.<sup>17</sup> Interestingly, in the model, the PDCF would not be tapped by investors in equilibrium, even though it is essential for ruling out the possibility of coordination failures and, in this way, stabilizing the market.

Martin, Skeie, and von Thadden (2010) (see, also, Copeland, Martin, and Walker [2010]) consider the game played by investors in the case when there is no "morning unwind." In the context of their model, they show that the investors' game is no longer a coordination game and, hence, runs can no longer happen. Their model is different, yet related to the model presented here. In particular, they consider the case where there are old and new investors playing the game. Then, the result relies on the assumption that, without the unwind, the dealer gets to observe whether or not it will fail before making any payments to existing (old) investors. This removes the incentives of existing investors to run, even if no new investor is willing to fund the dealer. But, when existing investors do not run, the dealer can withstand a run by new investors, which removes the incentives for new investors to run.

One way to obtain a similar result in our setup is by assuming that, barring daylight credit from the clearing bank, the dealer needs to arrange repo funding before making any investments. Also, let us assume that the dealer goes ahead with the investment only if it is able to convince both investors to fund the operation. In this situation, the payoff to an investor that agrees to enter a repo

<sup>&</sup>lt;sup>17</sup> This role of the PDCF is highlighted by Adrian, Burke, and McAndrews (2009) when they say: "The PDCF has the potential to benefit trading in the repo market beyond the direct injection of funding. The very existence of the facility is a source of reassurance to the primary dealers and their customers." Dudley (2009) also says that "the PDCF essentially placed the Fed in the role of the tri-party repo investor of last resort thereby significantly reducing the risk to the clearing banks that they might be stuck with the collateral. As a consequence, the PDCF reassured end investors that they could safely keep investing. This, in turn, significantly reduced the risk that a dealer would not be able to obtain short-term funding through the tri-party repo system."

contract, when the other investor does not, is the same as the payoff from not entering a repo contract; i.e., it is equal to zero. Assuming, as Martin, Skeie, and von Thadden (2010) do, that in case of indifference an investor agrees to repo, we have that the "unique" equilibrium in the investors' game is to play (R, R), and the dealer always gets funded.

#### **Correlated Equilibrium**

In the SPN equilibria of Proposition 4, the clearing bank in the morning has no doubts about the events that will take place during the afternoon when the cash investors have to decide whether or not to fund the securities dealer: Either the bank anticipates that funding from cash investors will be broadly available or it anticipates that no investor will be accepting repo requests. In principle, however, the bank may not be sure about the availability of funding in the afternoon. A simple representation of this uncertainty can be accomplished by using the alternative equilibrium concept of *correlated equilibrium*. <sup>18</sup>

In particular, suppose that at the time when investors have to decide whether or not to fund the dealer in the afternoon of the first date, they observe a public signal that can take two possible values:  $\alpha$  with probability  $\pi$ , and  $\beta$  with probability  $1 - \pi$ . Suppose also that, when investors observe  $\alpha$ , they play the equilibrium with actions (R, R), and when they observe  $\beta$ , they play the equilibrium with actions (NR,NR). The bank, instead, does not observe the public signal at the time when it has to decide whether or not to allow the dealer to incur an overdraft on its account at the bank.

**Proposition 5** Define  $\overline{\pi} \equiv y_B/(x_B + y_B)$ . If  $\pi \geq \overline{\pi}$ , then there is a correlated equilibrium in which the bank plays O in node 1 of Game 3. If  $\pi < \overline{\pi}$ , then there is a correlated equilibrium in which the bank plays NO in node 1.

The proof of the proposition is very similar to the other proofs and is not included here. We can interpret  $\pi$  as the clearing bank's perception of the likelihood that the dealer will obtain funding in the afternoon. If the probability is high enough, above the threshold  $\overline{\pi}$ , then the bank will agree to grant an overdraft. Note that, after the bank allows for the overdraft, with probability  $1-\pi$ , investors do not agree to fund the dealer in the afternoon and the clearing bank is stuck with the securities that served as collateral for the overdraft. In such case, the bank suffers a loss given by  $y_B$ . Note that, as the loss increases, the threshold value  $\overline{\pi}$  increases and gets closer to unity.

<sup>&</sup>lt;sup>18</sup> There is also a mixed-strategy SPN equilibrium of Game 3 in which investors randomize over actions R and NR, playing R with probability  $z_I/(x_I+z_I)$ . In such an equilibrium, the bank also faces uncertainty about the ability of the dealer to get funding at the end of period 1. However, we find the interpretation of this equilibrium less appealing and, for this reason, we do not discuss it here.

In other words, as the loss for the clearing bank becomes larger, the bank needs to be more and more certain that investors will fund the dealer in the afternoon if an overdraft is to be granted in the morning. We can think that a lower  $\pi$  represents a situation where confidence in the ability of the dealer to participate in the repo market decreases. If the situation deteriorates enough, to the point when  $\pi$  gets below the threshold  $\overline{\pi}$ , then the clearing bank will not agree to grant an overdraft (or unwind previously arranged repo transactions by granting the dealer daylight credit).

Note that, in contrast to the situation described in the previous subsection, here the payoff of the bank in case the dealer defaults after the morning unwind is relevant for the outcomes of the game. In the equilibrium of Proposition 5, the clearing bank retains some uncertainty about the ability of the dealer to obtain repo funding in the afternoon of date 1. The key to this result is that the public signal is only observed after the morning unwind and, hence, it creates the potential for a sudden shift in the behavior of investors in the afternoon repo market. Coordination failures are, perhaps, more likely to happen abruptly since they are based only on changes in the beliefs of market participants about the behavior of other market participants. Instead, changes in behavior driven by fundamentals, such as the ones studied in Propositions 2 and 3, are more likely to happen gradually over time, allowing the clearing bank to potentially exploit its informational advantage.

For concreteness, we have considered here a situation with only two investors. However, in general, there could be many more cash investors. <sup>19</sup> An alternative formalization would be to have a continuum of investors deciding at the end of date 1 whether or not to fund the dealer via repo transactions. In such case, it is clear that the decision of any one investor will not have a material consequence on the overall ability of the dealer to fund itself. In other words, if an investor enters a repo contract with a dealer when all the other investors do not, then the dealer will indeed fail and the investor with the repo contract will be stuck with the securities. The structure of payoffs that implies a coordination game arises more naturally in this case, relative to the case where there are only two investors. However, given our assumptions on payoffs, the results would be basically the same in both cases.

# 2. DISCUSSION

From the perspective of cash investors, the tri-party repo contract is almost equivalent to an interest-bearing demand deposit. Because of the daily unwind, investors have access to their cash during the day (on demand). During the

<sup>&</sup>lt;sup>19</sup> Copeland, Martin, and Walker (2010) consider a coordination game similar to the one studied here but where there are three investors in the game. See also Martin, Skeie, and von Thadden (2010).

night, the cash is locked in with the repo transaction. The next morning, the contract entitles the investor to a positive interest payment. In an uninsured demand deposit contract, investors are exposed to counterparty credit risk. In contrast, the tri-party repo contract could be considered, in principle, a form of secured lending since there is collateral pledged to address default risk. Haircuts on the collateral could be set so as to leave the lender with virtually no exposure to credit risk. However, in reality, evidence suggests that cash investors still perceive themselves as being exposed to some risk of losses when the borrower defaults (see, for example, Copeland, Martin, and Walker [2010] and PRC Task Force [2010]). We have taken the possibility of losses as a premise for our model, without trying to explain the fundamental reasons for under-collateralization. Understanding how this arrangement could arise optimally is not an easy task. Lacker (2001) provides a framework to think about collateralized debt that could be used to address these kinds of issues (see, also, Dang, Gorton, and Holmström [2010]). More work is clearly needed in this area.

In the United States, paying interest on demand deposits was not allowed until very recently. This restriction was especially binding for businesses. However, the financial system has developed some alternatives that constitute close substitutes of interest-bearing demand deposits. The tri-party repo arrangement could be considered one such alternative. The newly enacted Dodd-Frank financial reform legislation includes a provision that repeals the prohibition of paying interest on demand deposits and, starting on July 21, 2011, banks are now allowed to pay interest on these accounts. It is an open question how this will impact the tri-party repo market in the long run. It seems plausible that some cash investors looking for a way to earn interest on their cash holdings overnight may now turn to demand deposits at banks for this purpose. But, of course, there is a demand as well as a supply side in the tri-party repo market. On the demand side, securities dealers will still need to fund their portfolios of securities. Some form of repo contract is likely to play a role in satisfying that demand.

As we have explained, the source of funding for tri-party repos is two-fold: during the night, cash investors provide the funding and, during the day, daylight overdrafts granted by the clearing banks provide (most of) the funding. Some (if not most) of the cash owned by cash investors does not leave the books of the clearing bank during the day. Those funds are effectively demand deposits held by cash investors in their accounts at the clearing bank. These deposits, then, can be used by the clearing bank to fund the daylight credit provided to the dealers as part of the tri-party repo contract. But, to the extent that some of the cash owned by investors is used during the day to make payments and other transfers, the clearing bank needs to obtain daylight funding for the overdraft granted to the dealers. Of course, one readily available source of daylight funding for clearing banks is their daylight overdraft

capabilities with the Federal Reserve. If we think that the rate charged by the Fed for daylight credit is intentionally kept low ("to ensure the smooth functioning of payment and settlement systems"), then we could conclude that, to a certain extent, the tri-party repo arrangement is an indirect way for dealers to access subsidized funding during the day.<sup>20</sup>

With its simplified treatment of the events associated with a dealer's default, our formal analysis could not be used to address some significant issues being discussed in policy circles (see, for example, Copeland et al. [2011]). For example, the possibility that the liquidation of a dealer's portfolio could result in fire sale prices and externalities to other dealers (and to market participants in general) was left unexplored.<sup>21</sup> Another important issue that was not examined here is the possibility of "a loss of confidence" in the solvency of a clearing bank. This was a major concern for policymakers during the crisis and has been a salient point in the discussions about possible reforms to the infrastructure in the tri-party repo market (Bernanke 2008). Each clearing bank in the United States provides services to multiple dealers and to a large number of investors. To some extent, dealers need the clearing bank for their daily operation. It seems plausible, then, that problems at a clearing bank could spread to its client dealers if, for example, those dealers were relying on daylight credit to stay in business. Furthermore, cash investors usually have large unsecured exposures to their clearing bank during the day that could also destabilize them if that cash were no longer readily available. These are important issues that deserve careful consideration and are certainly related to the subject of this article. Here, however, we chose to keep the model simple on these dimensions to be able to sharpen our understanding of the strategic interaction between the clearing bank and investors, which may play a crucial role in the functioning of this complex market during a crisis.

In May 2010, the Tri-Party Repo Infrastructure Reform Task Force issued a set of recommendations to increase the stability of this market (PRC Task Force 2010). Their main proposal was to reform the system in order to reduce as much as possible the reliance of market participants on large amounts of intraday credit provided by clearing banks. In short, the proposal calls for an elimination of the indiscriminate daily unwind of all tri-party repo trades. Evidently, reducing the credit exposure of the clearing banks will limit the power of some of the strategic interactions highlighted in this article. However,

<sup>&</sup>lt;sup>20</sup> Currently, the Fed provides daylight credit to depository institutions using a two-tiered fee schedule. Those institutions that pledge enough acceptable collateral with their Reserve Bank receive daylight credit (up to a cap) at no charge. Uncollateralized daylight credit is charged a fee that is calculated per minute using an annual rate of 50 basis points. This system was only recently introduced. During the crisis, the Fed charged a uniform rate of 36 basis points for intraday credit and this credit was all uncollateralized. For more information on the current system see www.federalreserve.gov/paymentsystems/psr\_policy.htm.

<sup>&</sup>lt;sup>21</sup> For a model that is useful to address some of these issues, see Acharya and Viswanathan (2011).

if the morning unwind creates some valuable operational advantages that make the tri-party repo contract especially attractive to dealers and investors, then an obvious tradeoff arises between stability and effectiveness. <sup>22</sup> In such a case, fragility is not to be combated at all costs. As in many other situations where a risk-return frontier results in a tradeoff, the optimal arrangement could very well involve actually tolerating some positive risk.

There are also other alternatives that have been considered to limit this source of fragility in the tri-party repo market. For example, a system of capital requirements and risk charges that penalizes the intraday exposure of the clearing banks may give the appropriate incentives to participants, inducing them to move away from their over-reliance on intraday credit from the clearing banks (Tuckman 2010). Similarly, changes in the treatment of repos under bankruptcy law, such as removing them from the exception to the automatic stay (Roe 2009), could make these contracts less attractive and, hence, reduce the size of this potentially destabilizing market.

As the process of evaluating possible reforms continues, it is important to keep in mind that many of the features of the tri-party repo contract that we observe in the data are contingent on a set of rules (and common practices) that existed when the data was collected. If some of those rules are changed (by fiat or by newly built consensus among major participants), then some prevalent characteristics of the existing contract may also change. A case in point is the distribution of maturity terms in the market. Currently, term trades represent 10 percent to 40 percent of the market (PRC Task Force 2010). To the extent that participants stop perceiving the morning unwind as an automatic event for repos of longer maturities, it seems plausible that an even higher proportion of the outstanding repos will become overnight contracts. This may seem a fairly obvious point, yet it clearly highlights the limitations of evaluating the effects of possible changes in policies using only historical data. To complement our data analysis, we need to develop better models of the tri-party repo market that can allow us to conduct policy evaluations in a more meaningful way. The alternative is a costly process of trial and error purely based on experience in the actual market. Considering the current importance of this market, pushing forward a model-based agenda for studying this market seems worthwhile. The model introduced in this article is an attempt to take a preliminary step in this direction.

<sup>&</sup>lt;sup>22</sup> For example, changing to a system in which repos get unwound only later in the day (or, not unwound at all, in the case of term repos) will make those contracts less comparable with a demand deposit from the perspective of cash investors. While it is true that during the day investors are unlikely to need all the cash used in tri-party repos, the option to have that cash available presumably has some value for investors.

#### 3. CONCLUSION

In this article, we study a simple model of the strategic interaction between investors and the clearing bank in the tri-party repo market. In order to be able to apply simple game theory techniques to the problem, we abstract from many important features of this complex market. We mention several of them along the way in the presentation. Clearly, a lot more work is needed to extend the formal analysis in ways that would allow us to evaluate the role, and the relative importance, of those various features left unexplored here.

Perhaps the aspect most clearly highlighted by the model in this article is the role in the inception of a crisis played by participants' anticipation of each others' perceptions and actions. In particular, the model eloquently illustrates how changes in expectations about future events and actions can make a crisis happen abruptly before the fundamental factors behind it visibly manifest themselves. We conclude, then, that swings in perceptions (about fundamentals or about market confidence) can, in principle, trigger sudden crises in the tri-party repo market.

#### REFERENCES

- Acharya, Viral V., and S. Viswanathan. 2011. "Leverage, Moral Hazard, and Liquidity." *Journal of Finance* 66 (February): 99–138.
- Adrian, Tobias, Christopher R. Burke, and James J. McAndrews. 2009. "The Federal Reserve's Primary Dealer Credit Facility." Federal Reserve Bank of New York *Current Issues in Economics and Finance* 15 (August).
- Armantier, Olivier, Eric Ghysels, Asani Sarkar, and Jeffrey Shrader. 2011. "Stigma in Financial Markets: Evidence from Liquidity Auctions and Discount Window Borrowing during the Crisis." Federal Reserve Bank of New York Staff Report 483 (January).
- Bernanke, Ben S. 2008. "Financial Regulation and Financial Stability." Speech at the FDIC Forum on Mortgage Lending for Low and Moderate Income Households. Arlington, Va., July 8.
- Copeland, Adam, Antoine Martin, and Michael Walker. 2010. "The Tri-Party Repo Market before the 2010 Reforms." Federal Reserve Bank of New York Staff Report 477 (November).
- Copeland, Adam, Darrell Duffie, Antoine Martin, and Susan McLaughlin. 2011. "Policy Issues in the Design of Tri-Party Repo Markets." Unpublished manuscript, Federal Reserve Bank of New York.

- Dang, Tri Vi, Gary Gorton, and Bengt Holmström. 2010. "Financial Crises and the Optimality of Debt for Liquidity Provision." Unpublished manuscript, Yale University.
- Dudley, William C. 2009. "More Lessons from the Crisis." Remarks at the Center for Economic Policy Studies (CEPS) Symposium. Princeton, N. J., November 13.
- Duffie, Darrell. 2010. "The Failure Mechanics of Dealer Banks." *Journal of Economic Perspectives* 24 (Winter): 51–72.
- Ennis, Huberto M., and John A. Weinberg. 2010. "Over-the-Counter Loans, Adverse Selection, and Stigma in the Interbank Market." Federal Reserve Bank of Richmond Working Paper 10-07 (April).
- Federal Reserve Bank of New York. 2010. "Tri-Party Repo Infrastructure Reform." Federal Reserve Bank of New York White Paper (May 17).
- Gorton, Gary, and Andrew Metrick. 2010. "Haircuts." Federal Reserve Bank of St. Louis *Review 92* (November/December): 507–19.
- Gorton, Gary B., and Andrew Metrick. Forthcoming. "Securitized Banking and the Run on Repo." *Journal of Financial Economics*.
- Lacker, Jeffrey M. 2001. "Collateralized Debt as the Optimal Contract." *Review of Economic Dynamics* 4 (October): 842–59.
- Lubben, Stephen J. 2010. "The Bankruptcy Code Without Safe Harbors." *American Bankruptcy Law Journal* 84 (Spring): 123.
- Martin, Antoine, David Skeie, and Ernst-Ludwig von Thadden. 2010. "Repo Runs." Federal Reserve Bank of New York Staff Report 444. (April).
- Osborne, Martin J., and Ariel Rubinstein. 1994. *A Course in Game Theory*. Cambridge, Mass.: The MIT Press.
- PRC Task Force. 2010. Report of the Payments Risk Committee Task Force on Tri-party Repo Infrastructure. (May 17)
- Roe, Mark. 2009. "End Bankruptcy Priority for Derivatives, Repos and Swaps." *Financial Times*, December 16.
- Tuckman, Bruce. 2010. "Systemic Risk and the Tri-Party Repo Clearing Banks." CFS Policy Paper (February 2).

# K-Core Inflation

Alexander L. Wolman

Standard measures of inflation (for example, personal consumption expenditure [PCE] or consumer price index [CPI]) are constructed in order to accurately describe the behavior of consumption prices as a whole. However, to the extent that the inflation rate in a given period is accounted for by large relative price changes for particular goods and services, it may be desirable to have additional measures of inflation that adjust for those large relative price changes. These alternatives would be useful if large relative price changes are a source of noise, obscuring underlying patterns in the economy. Any such alternative inflation measure could never be the best measure of overall price changes, but it might provide valuable information about the behavior of future inflation, or more generally about the "state of the world" relevant for conducting monetary policy. This article describes a new class of measures of underlying inflation called "k-core inflation."

The term "core inflation" came into use in the 1970s, when large price increases for food and energy coincided with high overall CPI inflation and, in some years, with weak economic activity. Researchers using a Phillips curve framework at that time sought a notion of inflation that was consistent with a positive association between inflation and real activity. For example, Robert Gordon (1975) referred to "underlying 'hard-core' inflation" as distinct from the contributions made by food and energy, dollar devaluation, and the end of price and wage controls. The Bureau of Labor Statistics responded to these conditions in 1977 by beginning to publish a measure of the CPI that omitted food and energy components ("the index for all items less food and energy"). Today that subindex of the CPI is widely referred to as the core CPI, and, more

The author thanks Todd Clark, Marianna Kudlyak, Nika Lazaryan, Thomas Lubik, and Ned Prescott for helpful comments. The views in this paper are the author's and do not represent the views of the Federal Bank of Richmond, the Federal Reserve Board of Governors, or the Federal Reserve System. E-mail: alexander.wolman@rich.frb.org.

<sup>&</sup>lt;sup>1</sup> Beryl Sprinkel (1975) seems to have used the same term ("hard-core" inflation) a few months earlier than Gordon.

<sup>&</sup>lt;sup>2</sup> See the Bureau of Labor Statistics (2011).

generally, "core inflation" is understood to refer to some broad price index that excludes food and energy contributions.

Although both the term "core inflation" and the CPI measure originated in the 1970s, it was not until around 1990 (see Ball and Cecchetti [1990]) that the two became essentially synonymous. In recent years, economists have proposed many alternative measures of core inflation. One of the more prominent alternatives is trimmed mean inflation, which removes from the inflation calculation those price changes above and below specified percentiles in the distribution. K-core inflation, the measure proposed in this article, is a close cousin both to the standard core inflation measure (the index for all items less food and energy) and to trimmed mean inflation. Instead of removing food and energy prices—as core does—and instead of removing prices beyond specified percentiles in the distribution—as a trimmed mean does—k-core inflation removes items whose relative prices change by more than a specified threshold. If one's objective is to construct a measure of inflation on which large relative price changes have a limited effect, then k-core inflation seems clearly preferable to both inflation ex-food and energy and trimmed mean inflation. In periods during which food and energy prices move with the overall price level, whereas other categories experience large relative price changes, inflation exfood and energy will exclude small relative price changes and include large relative price changes. In contrast, k-core inflation will always exclude—and only exclude—the large relative price changes. Likewise, in periods during which the distribution of relative price changes is unusually concentrated but asymmetric, trimmed mean inflation would exclude many small relative price changes, and could produce a measure that deviates markedly from overall inflation. In contrast, k-core inflation would simply replicate overall PCE inflation if there were no large relative price changes.

Section 1 provides some background information on the construction of PCE inflation and the behavior of the category price changes that go into constructing PCE inflation. Section 2 describes k-core inflation in general terms. Whereas the measure in Section 2 is a parametric family, in Section 3 we show how the properties of k-core inflation vary with that parameter (k, the criterion for a large relative price change). We specify a value for k and compare k-core inflation to core inflation and trimmed mean inflation. Section 4 suggests areas for future research and concludes.

# 1. INFLATION AND THE DISTRIBUTION OF PRICE CHANGES

The two most commonly discussed measures of inflation in the United States are PCE inflation and CPI inflation. PCE inflation is an index of the rate of price change for a broad array of consumption goods and services—technically the entirety of consumption in the national income and product accounts. CPI inflation is an index of the rate of price change for "out-of-pocket"

consumption expenditures. As such, there are a number of differences between the components of PCE inflation and those of CPI inflation. Most importantly, PCE inflation puts a significantly higher weight on health care spending, and CPI inflation puts a significantly higher weight on housing services. There are also differences in the way the indexes are calculated; for details, see Clark (1999). Because the PCE index is a more comprehensive measure and is generally believed to be a more accurate measure of overall price changes, in the remainder of this article all references to inflation (without other qualifiers) will be to PCE inflation.

PCE inflation ( $\pi_t$ ) is a Fisher ideal index of price changes for a large number (N) of categories of consumption goods; it is the geometric mean of the Paasche and Laspeyres indexes of price change. The Paasche index in period t, denoted  $\pi_t^P$ , is the rate of price change from period t - 1 to period t for the consumption basket purchased in period t:

$$\pi_t^P = \frac{\sum_{n=1}^N p_{n,t} q_{n,t}}{\sum_{n=1}^N p_{n,t-1} q_{n,t}}.$$
 (1)

In this expression,  $p_{n,t}$  and  $q_{n,t}$  are the price and quantity purchased of category n in period t. The Laspeyres index in period t, denoted  $\pi_t^L$ , is the rate of price change from period t-1 to period t of the consumption basket purchased in period t-1:

$$\pi_t^L = \frac{\sum_{n=1}^N p_{n,t} q_{n,t-1}}{\sum_{n=1}^N p_{n,t-1} q_{n,t-1}}.$$
 (2)

Thus, the PCE inflation rate is given by the following formula:

$$\pi_{t} = \sqrt{\left(\frac{\sum_{n=1}^{N} p_{n,t} q_{n,t}}{\sum_{n=1}^{N} p_{n,t-1} q_{n,t}}\right) \left(\frac{\sum_{n=1}^{N} p_{n,t} q_{n,t-1}}{\sum_{n=1}^{N} p_{n,t-1} q_{n,t-1}}\right)}.$$
 (3)

Note that both the Paasche and Laspeyres indexes can be written as weighted averages of price changes for each category, thus

$$\pi_{t} = \sqrt{\left(\sum_{n=1}^{N} \omega_{n,t-1}^{P} \pi_{n,t}\right) \left(\sum_{n=1}^{N} \omega_{n,t-1}^{L} \pi_{n,t}\right)},$$
(4)

where  $\pi_{n,t}$  is the rate of price change for consumption category n in period t (that is,  $\pi_{n,t} = p_{n,t}/p_{n,t-1}$ ), and where the weights are given by

$$\omega_{n,t-1}^{L} = \frac{p_{n,t-1}q_{n,t-1}}{\sum_{j=1}^{N} p_{j,t-1}q_{j,t-1}}$$
 (5)

and

$$\omega_{n,t-1}^{P} = \frac{p_{n,t-1}q_{n,t}}{\sum_{i=1}^{N} p_{i,t-1}q_{i,t}}.$$
 (6)

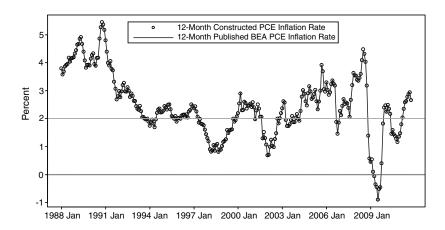


Figure 1 PCE Inflation: Constructed and from the BEA

The Laspeyres weight,  $\omega_{n,t-1}^L$ , is the period t-1 expenditure share for category n, and the Paasche weight,  $\omega_{n,t-1}^P$ , is the hypothetical expenditure share associated with evaluating the period t consumption basket at period t-1 prices.

Hundreds of consumption categories comprise PCE inflation, which is compiled by the Bureau of Economic Analysis of the U.S. Department of Commerce (BEA). We aggregate some of those categories in order to have a consistent panel going back to January 1987, and are left with 240 categories, covering 100 percent of personal consumption expenditure, for the period from January 1987–October 2011. Figure 1 plots the behavior of official 12-month PCE inflation over this period (solid line), together with the series we constructed using (4) with 240 categories (open circles). A careful look at the figure reveals slight differences between the two measures in some periods. Overall however, the two series are close enough that it appears legitimate to proceed using the constructed PCE measure instead of the BEA's measure.

If the component price changes that aggregate up to PCE inflation were all close to each other, and thus close to PCE inflation, then there would be no reason to consider inflation measures that control for large relative price changes. The black line in Figure 2 displays the distribution of relative price changes for all categories across all periods in the sample, where the relative price change for category n in period t is simply the difference between the rate of price change for that category and the rate of PCE price change:

$$r_{n,t} = \pi_{n,t} - \pi_t.$$

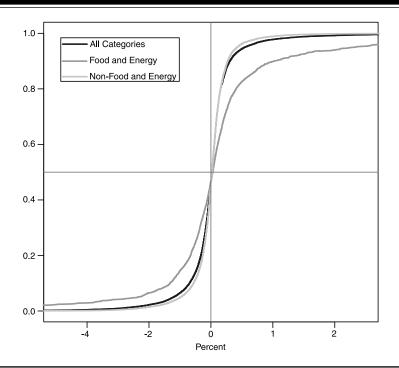


Figure 2 CDF of Monthly Relative Price Changes, Jan. 1987–Oct. 2011

To construct the distribution, each  $r_{n,t}$  is weighted by the corresponding expenditure share  $\omega_{n,t}^L$ . The distribution of monthly relative price changes is indeed concentrated around zero, with an interquartile range of (-0.23 percent, 0.25 percent). However, there are also many large relative price changes: For example, 12.1 percent of (weighted) relative price changes are greater than 1 percent per month in absolute value. Figure 2 also displays the distribution of relative price changes for the 28 food and energy categories (dark gray) and for the 212 non-food and energy categories (light gray). Food and energy relative prices vary much more than their complement: The interquartile range for food and energy categories is (-0.53 percent, 0.55 percent) compared to (-0.19 percent, 0.24 percent) for other categories.

In sum, from Figure 2 it is clear that (i) there is nontrivial variation in the relative prices of different categories of consumption, and (ii) the variation is especially large for food and energy categories. We take those facts as motivation for constructing measures of inflation that attempt to control for

the contributions of *large* relative price changes.<sup>3</sup> We refer to any such measure below as a measure of underlying inflation.

#### 2. OLD AND NEW MEASURES OF UNDERLYING INFLATION

Because food and energy prices are so much more volatile than the prices of other consumption categories (see Figure 2), a natural underlying inflation measure is one that simply removes food and energy prices from the inflation calculation. This measure, so-called "ex-food and energy" PCE inflation, has the virtue of simplicity. However, always and only removing food and energy prices does not mean always and only removing categories with the largest relative price changes. Of the top 10 price increases and the top 10 price decreases each period, on average less than one quarter of those largest price changes were from food and energy categories. And of the 20 smallest relative price changes each period (measured by absolute value), more than 8 percent were from food and energy categories.<sup>4</sup> Thus, removing only food and energy price changes means not removing most of the large relative price changes, and it means removing a significant number of very small relative price changes.

An alternative to ex-food and energy inflation that does remove only the largest price changes each period is trimmed mean inflation. Trimmed mean inflation begins with the weighted cumulative distribution function (CDF) of monthly price changes each period, and removes those price changes that lie outside upper and lower percentile cutoffs. If the upper and lower cutoffs are the 50th percentile, then trimmed mean inflation is simply the rate of price change for the median category. Bryan and Cecchetti (1994) and Dolmas (2005) provide detailed discussions of trimmed mean inflation, with the former focusing on the CPI and the latter on PCE inflation. They suggest various methods of choosing the specific percentile cutoffs for trimmed mean inflation. The Federal Reserve Bank of Dallas maintains a trimmed mean inflation series (Federal Reserve Bank of Dallas 2012)—currently, their preferred cutoffs are 24 percent from the bottom of the distribution and 31 percent from the top (see Section 3 for further discussion). From the data behind Figure 2, on average these criteria remove relative price decreases greater than 0.25 percent per month, and relative price *increases* greater than 0.18 percent per month.

If the goal is to construct a measure of underlying inflation by removing large relative price changes, then a trimmed mean has an obvious advantage

<sup>&</sup>lt;sup>3</sup> From the definition of PCE inflation, it is tautological that *all* relative prices together account completely for the behavior of inflation.

<sup>&</sup>lt;sup>4</sup> These statements refer to unweighted price changes—meaning that each category is weighted equally. There are 28 food and energy categories out of 240 total categories in our sample, so that if price change distributions were identical across categories then 11.6 percent of any range of price changes would be from food and energy categories.

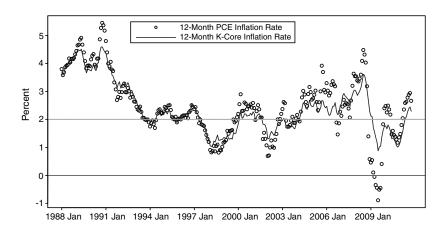
6 Max 5 4 75th Percentile 3 Mean Median -2 25th Percentile Min 0.00 0.05 0.10 0.15 0.20 k

Figure 3 How the Distribution (Across Time) of k-Core Inflation Varies with k

relative to ex-food and energy inflation: It removes categories with the largest price changes, regardless of whether they are food and energy categories. However, the fact that a trimmed mean removes fixed percentiles of each period's distribution of price changes has an important implication. Depending on how the distribution of price changes behaves in a given period, price changes of different sizes will be removed. That is, once one specifies the percentile cutoffs, the largest price changes are removed each period, regardless of the size of those price changes. But if the goal is to remove large relative price changes that will be removed and hold that size fixed each period. In the remainder of the article we consider such a measure, which we call k-core inflation.<sup>5</sup>

K-core inflation specifies a cut-off value, k, for the size of relative price changes. If the relative price change for category n is less than k in absolute

<sup>&</sup>lt;sup>5</sup> Researchers such as Bryan and Cechetti (1994) and Dolmas (2005) motivate trimmed mean inflation partly on statistical grounds and partly on theoretical grounds. In the conclusion, we suggest a theoretical grounding for soft-core inflation.



**Figure 4 Benchmark k-Core Inflation (k = 2 Percent Annual Rate)** 

value, then the price change for category n is included without modification. If the relative price change for category n in period t is greater than k in absolute value, then the price change for category n is truncated at k. Formally, for a given k, k-core inflation  $(\pi_t^{sc})$  is defined as follows:

$$\pi_{t}^{kc}(k) = \sqrt{\left(\sum_{n=1}^{N} \omega_{n,t-1}^{P} \pi_{n,t}^{kc}(k)\right) \left(\sum_{n=1}^{N} \omega_{n,t-1}^{L} \pi_{n,t}^{kc}(k)\right)},$$
 (7)

where

$$\pi_{n,t}^{kc}(k) = \begin{cases} \pi_{n,t}, & \text{if } |\pi_{n,t} - \pi_t| < k \\ \pi_t \left( 1 + k \cdot sign\left( \pi_{n,t} - \pi_t \right) \right), & \text{if } |\pi_{n,t} - \pi_t| \ge k \end{cases}$$
 (8)

Three assumptions embodied in this definition require some discussion. First, large price changes are truncated rather than being omitted. This choice is based on the facts that there is uncertainty about the proper value of k, and about whether or not every relative price change greater than k should be omitted from underlying inflation. An appealing implication of this assumption is that varying k between zero and infinity makes  $\pi_t^{kc}(k)$  a smooth function that starts and ends at  $\pi_t$ . For low k all price changes are replaced with actual inflation, and for high k all price changes are admitted, which returns actual inflation. The second important assumption is that the criterion for truncating price changes is the size of relative price change, rather than the size of nominal price change. This choice simply reflects the view that it is large *relative* price changes that we want to control for. Third, the criterion (k) does not vary with the level of inflation. There is a large literature on the relationship between

	Mean	Min	25th and	75th Percentiles	Max	Std. Dev.
PCE	2.51	-13.55	1.40	3.74	13.66	2.41
k-core	2.43	-1.31	1.43	3.37	6.28	1.38
XFE	2.56	-5.35	1.59	3.37	8.30	1.55
Trimmed Mean	2.47	0.30	1.87	2.99	5.68	0.95

Table 1 Summary Statistics for One-Month Inflation (Annualized Percent)

relative price variability and inflation (see Hartman [1991], for example). Based on that literature, one might argue that *k* should be an increasing function of the inflation rate. Because the data used in this article is from a period of relatively low and stable inflation, we assume that such considerations are not quantitatively important.

From Figure 2, one can see how the choice of k maps into the fraction of price changes that will be truncated:  $k \ge 0.04$  (4 percent monthly in Figure 2) would mean truncating a tiny fraction of price changes, whereas k = 0.005 (one-half percent monthly) would mean truncating 13.9 percent of weighted price changes because of relative price decreases, and 12.7 percent because of relative price increases. Of course, these are averages, and the fraction of expenditures (equivalently, price changes) affected in a given period would depend on the distribution of price changes in that period.

#### 3. BEHAVIOR OF K-CORE INFLATION

Figure 3 plots summary statistics for 12-month k-core inflation as a function of k, using the entire sample. For each value of k, we compute the time series for k-core inflation and plot the summary statistics, mean, median, maximum, minimum, and 25th and 75th percentiles. The figure shows how these summary statistics of the time series vary with k. For low and high values of k, the statistics are similar, reflecting the fact that k-core inflation converges to overall PCE inflation as k approaches zero or infinity. The properties of k-core inflation are sensitive to k for values around 0.02 (2 percent monthly relative price change). The range (maximum minus minimum) of k-core inflation shrinks from almost six percentage points (the range for PCE inflation) for high and low k to less than four percentage points when k is around 0.02. Because it is a round number and comes close to minimizing the range of k-core inflation, we will use k = 0.02 as our benchmark.

Figure 4 plots the time series for benchmark k-core inflation, together with overall PCE inflation (the constructed measure from Figure 1). Although we construct k-core inflation as a monthly measure, using (7), the time series plotted in Figure 4 and in subsequent figures display the 12-month cumulative

Mean Min 25th and 75th Percentiles Max Std. Dev. PCE 2.46 -0.911.86 3.02 5.42 1.11 4.59 k-core 2.41 0.88 1.81 2.88 0.88 **XFE** 2.55 0.98 1.84 2.67 5.19 1.10 Trimmed Mean 2.45 2.05 2.66 0.80 4.28 0.74

**Table 2 Summary Statistics for 12-Month Inflation (Percent)** 

k-core inflation rate.<sup>6</sup> As expected from Figure 3, k-core inflation is notably less volatile than PCE inflation. The behavior of inflation in the depths of the Great Recession illustrates this point well: From mid-2008 to mid-2009, PCE inflation fell by more than five percentage points, whereas k-core inflation fell by less than three percentage points. However, it is not always the case that k-core inflation is a smoother version of PCE inflation. For example, in the second half of 2010, PCE inflation was relatively low (generally below 1.5 percent), yet k-core inflation was below PCE inflation.

# K-Core Inflation and Ex-Food and Energy Inflation

Having motivated k-core inflation as an appealing alternative to ex-food and energy inflation and trimmed mean inflation, we now compare the behavior of k-core to inflation ex-food and energy (henceforth XFE), and in the next section, to trimmed mean inflation (henceforth TMI). The top three rows of Tables 1 and 2 display summary statistics for one-month and 12-month PCE inflation, k-core inflation, and XFE.<sup>7</sup> For monthly price changes, both k-core and XFE are much less volatile than PCE inflation. This statement holds whether one measures volatility by max-min, standard deviation, or interquartile range. K-core inflation is less volatile than XFE, apart from the interquartile range measure. Moving from one-month to 12-month inflation, the comparisons become more muddied. Because each of these series has a substantial transitory component, the volatility of 12-month inflation is lower in every case than the volatility of the one-month measure. The transitory component is strongest in PCE inflation, so that the standard deviation of that series is cut by more than half when comparing one-month and 12-month changes. In contrast, the standard deviation of XFE inflation falls by just 29 percent, leaving the standard deviations of 12-month PCE and XFE inflation essentially identical. K-core's standard deviation is 36 percent lower for 12-month than one-month changes,

<sup>&</sup>lt;sup>6</sup> The only reason for doing this is that one-month inflation is quite volatile. Some of the tables, as well as Figure 2, refer to one-month price changes.

<sup>&</sup>lt;sup>7</sup> Note that the version of XFE analyzed here is not the version reported by the BEA. Instead, we calculate our own version by removing the 28 food and energy categories (and adjusting the other weights accordingly) in equation (3). The resulting series is close to the one reported by the BEA.

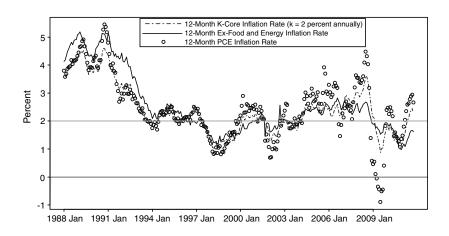


Figure 5 K-Core Inflation and Ex-Food and Energy Inflation

leaving it well below XFE (0.88 versus 1.10). However, the interquartile range for 12-month k-core inflation is well *above* that for XFE.

Figure 5 plots the time series for 12-month k-core inflation and XFE. Although Tables 1 and 2 indicate that k-core inflation is generally less volatile than XFE, Figure 5 shows that this volatility ranking is heavily influenced by the first few years of the sample, when PCE inflation was often above 5 percent. During that time, k-core inflation was well below XFE. In the last several years, by contrast, XFE has been markedly less volatile than k-core. The recent period has involved dramatic swings in energy prices. In September 2008 for example, 12-month inflation was 4.03 percent, whereas XFE was 2.52 percent. During this period, Figure 6 shows that there were many large relative price decreases that k-core inflation adjusted for, whereas XFE did not. As a result, k-core inflation was much higher than XFE, 3.37 percent, in the 12 months preceding September 2008.

From Figure 2, it is already clear that k-core inflation with k=0.02 does not always truncate food and inflation categories, and sometimes truncates categories other than food and inflation. Table 3 lists the 15 categories whose price changes are truncated most frequently when k=0.02, restricting to categories representing more than 0.01 percent of expenditure on average over the sample period.<sup>8</sup> Seven of the 15 categories are either food or energy categories (they are indicated in bold in the table). The 15 categories together

 $<sup>^{8}</sup>$  The restriction based on expenditure shares meant that two categories were eliminated and replaced with other categories.

Table 3 Categories Whose Relative Price Changes Most Frequently Exceed k = 0.02 in Absolute Value

Category	Freq. Exceed k	Avg. Share
Eggs	81	0.0007
Fuel Oil	80	0.0023
Gasoline & Other Motor Fuel	79	0.0269
Fresh Vegetables	64	0.0040
Indirect Securities Commissions	64	0.0013
Mutual Fund Sales Charges	63	0.0011
Air Transportation	52	0.0060
Direct Securities Commissions	44	0.0032
Used Truck Margin	42	0.0017
Natural Gas	41	0.0065
Fresh Fruit	39	0.0027
Other Fuels	39	0.0002
Luggage & Similar Personal Items	35	0.0024
Tobacco	32	0.0096
Commissions for Trust, Fiduciary, & Custody Activities	31	0.0011

Notes: Food and energy categories are listed in bold.

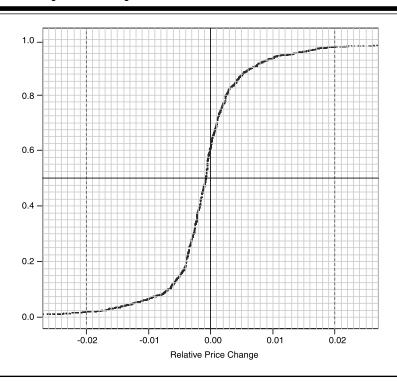
represent 7 percent of expenditures, and the seven food and energy categories represent 4.3 percent of expenditures.

#### K-Core Inflation and Trimmed Mean Inflation

Next, we compare our k-core inflation measure to TMI. To generate TMI we use a lower cutoff of 20 percent of expenditures, and an upper cutoff of 23 percent. Dolmas (2005) proposes three different criteria for choosing the upper and lower cutoffs. One of the criteria he uses is to minimize the squared deviations from a centered 36-month moving average of overall PCE inflation. Applying that criterion to our sample generates the 20 percent and 23 percent cutoffs. Note that for k-core inflation, our choice of k=0.02 nearly represents the value of k that would minimize the deviation of k-core inflation from the 36-month moving average of overall PCE inflation; that value is  $\tilde{k}=0.018$ . However, even this "optimized" version of k-core is considerably less successful than the optimized TMI at matching the moving average. The sum of squared deviations for the TMI is  $7.5 \times 10^{-5}$ , whereas the sum of squared deviations for k-core inflation is  $1.2 \times 10^{-4}$ .

Tables 1 and 2 contain summary statistics for TMI, in the bottom row, and Figure 7 plots TMI along with k-core inflation and PCE inflation. TMI is less volatile than either XFE or k-core inflation. The difference is especially striking for one-month inflation, where the standard deviation of TMI is at least 30 percent lower than that of the other measures, and the difference between the maximum and minimum values is 5.4 percent for TMI, compared to 7.6

Figure 6 Pooled Distribution of Monthly Relative Price Changes from Sept. 2007–Sept. 2008



percent for k-core and 13.7 percent for XFE. The relative stability of TMI compared to k-core can be partly understood by referring back to Figure 2, the distribution of relative price changes. Although k-core inflation is not a trimmed mean, we can think of it as a "truncated mean," where the percentile cutoffs for truncation (at 0.02) change each period. From Figure 2, on average both the lower and upper cutoffs for truncation are close to 0.25 percent of expenditure-weighted price changes. Thus, TMI with cutoffs at about 20 percent results in a price index that differs much more dramatically from PCE inflation than does our k-core inflation measure. If we were to exclude categories instead of truncating their price changes, the resulting series would be precisely a trimmed mean with time-varying cutoffs. From the numbers reported in the previous section we know that relatively little trimming would be implied by k=0.02. Figure 8 displays the somewhat smoother series generated by eliminating categories with k=0.02 instead of truncating their price changes.

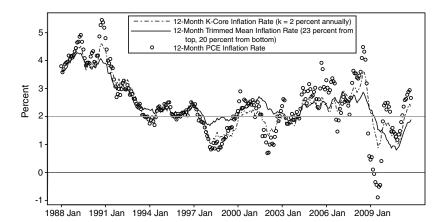


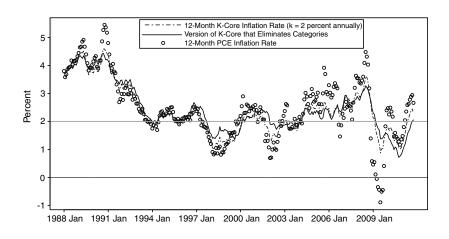
Figure 7 K-Core and Trimmed Mean Inflation

#### 4. CONCLUSION

We have proposed a new measure of underlying inflation, referred to as k-core inflation. All such measures are motivated to some degree by the idea that large relative price changes may represent noise, which the monetary authority ought to filter out for the purpose of forecasting or for inferring the current stance of monetary policy. K-core inflation does this filtering by specifying a threshold for a large relative price change. Relative price increases or decreases beyond that threshold are truncated to be equal to the threshold. In contrast, inflation ex-food and energy excludes food and energy prices regardless of how much those prices change, and trimmed mean inflation excludes fixed percentiles of the price change distribution, regardless of the size of price changes to which those percentiles correspond.

The figures and tables in the article illustrate how k-core inflation behaves, and how it compares to inflation ex-food and energy and to trimmed mean inflation. Mid-2008 was a period in which the differences between k-core inflation and these other measures were particularly large and persistent. K-core inflation indicated significantly higher underlying inflation in mid-2008 than either ex-food and energy inflation or trimmed mean inflation. The situation looks somewhat similar today, when energy price increases are again in the headlines: In the 12 months preceding October 2011, k-core inflation was

Figure 8 K-Core Inflation when Large Price Changes are Eliminated Rather than Truncated



2.3 percent, compared to 2.7 percent for overall PCE inflation, 1.6 percent for PCE inflation ex-food and energy, and 1.9 percent for trimmed mean inflation.

This article is exploratory in nature. It would be interesting to investigate k-core inflation further in at least three dimensions. First, measures of underlying inflation are often evaluated on the basis of their ability to forecast PCE inflation. How does k-core inflation fare according to this criterion? Second, the definition of k-core inflation used here has maintained that PCE inflation is the correct inflation rate against which to measure relative price changes. Perhaps k-core inflation is instead the correct inflation rate against which to measure relative price changes. Applying this change to our definition would require solving a fixed-point problem to compute k-core inflation. Finally, and most importantly, it would be interesting to pursue possible theoretical underpinnings of k-core inflation. If there are large sector-specific shocks (as suggested by much research on price adjustment, such as Golosov and Lucas [2007]) and if the structure of the economy and the behavior of monetary policy are such that monetary policy does not generate large relative price changes, then something like k-core inflation might be a useful indicator of monetary conditions. It would be straightforward to study this issue in a multisector equilibrium model. Of course, it is also possible that large relative price changes could actually signal loose monetary policy. That would go against the spirit of this article, but it cannot be ruled out a priori. Whether or not such a possibility is empirically relevant would seem to depend on the nature of

cross-sectoral variation in price stickiness and demand and supply elasticities. These issues could be studied in the context of a calibrated equilibrium model.

# **REFERENCES**

- Ball, Laurence, and Stephen G. Cecchetti. 1990. "Inflation and Uncertainty at Long and Short Horizons." *Brookings Papers on Economic Activity* 21 (1): 215–54.
- Bryan, Michael, and Stephen Cecchetti. 1994. "Measuring Core Inflation." In *Monetary Policy*, edited by N. Gregory Mankiw. Chicago: University of Chicago Press, 195–215.
- Bureau of Labor Statistics. 2011. "The So-called 'Core' Index: History and Uses of the Index for All Items Less Food and Energy." In *Focus on Prices and Spending*. Washington, D.C.: Bureau of Labor Statistics (February).
- Clark, Todd E. 1999. "A Comparison of the CPI and the PCE Price Index." Federal Reserve Bank of Kansas City *Economic Review* Q3: 15–29.
- Dolmas, Jim. 2005. "Trimmed Mean PCE Inflation." Federal Reserve Bank of Dallas Working Paper 0506 (July).
- Federal Reserve Bank of Dallas. 2012. "Trimmed Mean PCE Inflation Rate." Available at www.dallasfed.org/research/pce/index.cfm.
- Golosov, Mikhail, and Robert E. Lucas, Jr. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115: 171–99.
- Gordon, Robert J. 1975. "Alternative Responses of Policy to External Supply Shocks." *Brookings Papers on Economic Activity* 6 (1): 183–206.
- Hartman, Richard. 1991. "Relative Price Variability and Inflation." *Journal of Money, Credit and Banking* 23 (May): 185–205.
- Sprinkel, Beryl W. 1975. "1975: A Year of Recession, Recovery, and Decelerating Inflation." *Journal of Business* 48 (January): 1–4.

# The Cost of Unanticipated Household Financial Shocks: Two Examples

Kartik Athreya and Urvi Neelakantan

ouseholds sometimes experience unexpected negative changes to their financial circumstances. In this article, we quantify the consequences of two representative types of unanticipated financial shocks. By "unanticipated," we mean that households in our experiments are modeled as ignoring even the *possibility* that the shock could occur. We are thus interested in the cost of an event that comes as such a surprise to the household that its previous consumptions-savings decisions in no way prepared it for such an eventuality. Our analysis is therefore exactly analogous to a standard form of experiment in business cycle contexts, e.g., the impulse response of an unanticipated fiscal or monetary policy shock that agents know is permanent as soon as it occurs (see, for example, Baxter and King [1993]).

For each shock, our calculations tell us how costly it is for households to live in a world where the shock occurs compared to a world in which it does not. Why might such costs be useful to study? If the household (or a policymaker) could pay—say through investment in financial education—for information that would enable it to avoid the shock or mitigate its effect, our calculations may provide an upper bound on how much it might be willing to pay. The reason is that the cost of the shock depends not just on its magnitude but also its likelihood. For a shock of a given size, households will be less willing to pay to avoid it as its likelihood falls. We assume that the shock is

We thank Juan Carlos Hatchondo, Marianna Kudlyak, Ned Prescott, Pierre Sarte, Max Schmeiser, Jonathan Tompkins, John Weinberg, Kim Zeuli, seminar participants at the Federal Reserve Bank of Richmond and the University of Illinois at Urbana-Champaign, and participants at the Georgetown Center for Economic Research and the Southern Economics Association conferences for helpful comments and suggestions. We are solely responsible for any errors. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: kartik.athreya@rich.frb.org; urvi.neelakantan@rich.frb.org.

completely unanticipated or seen as one with zero likelihood. If its likelihood is truly close to zero, this makes it not worth doing much about, all else equal. Moreover, if the household is incorrect in assigning zero or near-zero likelihood to the shock, that is a belief that maximizes the amount by which households "underestimate" the risks. If the household instead knew that the shock could occur with positive probability, it would take actions (to the extent warranted by the magnitude and likelihood of the shocks, and the household's aversion to risk) to reduce its severity. By contrast, our model features households that will, by their unawareness, have made no provisions at all at the onset of either of the shocks we consider. Our cost calculations will also allow us to compare shocks, that is, to point out which shocks are costlier (assuming equal likelihood) and therefore worthy of greater attention.

The two types of shocks we consider here are (1) an unanticipated drop in net worth and (2) an unpredicted increase in borrowing costs for all forms of unsecured debt. Each is meant to represent the occurrence of an empirically plausible scenario. The first provides insight into the cost borne by those who are surprised by declines in the value of assets in their portfolio. Consider, for example, a household that has a net worth that is largely composed of equity in its home, and for which the recent decline in U.S. house prices came as a shock. It is evident that many commentators and experts placed little probability on a widespread decline in home prices. The second case is that of a sudden, widespread increase in the cost of rolling over debt and captures the effects of general credit market tightening as might occur in the midst of a severe recession that was a priori assigned zero probability. Note that both shocks are fully persistent.

The size of a shock is an inadequate measure of its importance to a household, in particular because the cost is likely to vary across households. Thus, quantifying the cost requires a model of household financial decision making. Households make consumption-savings decisions with the goal of smoothing consumption over their lifetime. A consequence of this hypothesis is that households' financial positions (and, hence, the cost of the shock to them) will differ by age. Moreover, to the extent that households face other, more predictable forms of risk throughout their lives, they will also differ from each other at any given age. In turn, the cost of a shock will vary across households of any given age as well. The economic model we use is a fairly standard version of a life-cycle model of consumption and savings, and follows Athreya (2008). We use the model and, in particular, the optimal value function of the household, to quantify the effects of the shocks. Specifically, we use the

<sup>&</sup>lt;sup>1</sup> Freddie Mac's "Rent or Buy" calculator provides anecdotal evidence of the lack of concern with house price declines. The calculator did not allow users to analyze the effects of negative realizations for home prices, even though the same device allowed for increases in house prices of up to 100 percent (Joffe-Walt and Davidson 2010).

model to determine the amount of annual consumption that a household would be willing to give up to avoid facing the shock.

The reader will no doubt see that our article is highly stylized. Importantly, it abstracts from portfolio choice and focuses instead on a simple scalar measure of net worth. In its current form it therefore cannot speak directly to particular kinds of financial decisions, such as house purchases or any other leveraged purchase of risky assets. In particular, our focus on net worth effectively precludes us from being able to assess the impact of decisions whose effects derive primarily from their impact on the gross financial positions of households—as well as on any attendant changes in the periodic payment obligations—while leaving net worth essentially unchanged. Our model also abstracts from the labor supply decision, which could mitigate the cost of the shock by allowing households to simply "work their way" out of a reduction in net worth. However, this is not wholly unreasonable because the shocks we consider are most relevant to recessionary settings, in which labor markets could plausibly preclude such adjustments.

Finally, our model embodies a strong assumption with respect to the information that households possess: We assume that the shocks that the household faces are completely unanticipated. It is possible, instead, that households are aware of the existence of the kinds of shocks we analyze in this article, but wrong about the exact probabilities with which they could occur. Nonetheless, while strong, this assumption allows us to determine what are likely to be upper bounds on the consequences of such shocks. After all, any information received in advance about the likelihood of such events can only make the eventual shock, if it occurs, easier to deal with, as households will have consumed and saved in anticipation of such possible outcomes. In addition, the current work is simply a small first step, and we have indeed begun to incorporate each of these features in ongoing work (Athreya, Ionescu, and Neelakantan 2011) that we hope will shed greater light on the questions addressed here.

With the preceding in mind, we describe the model in Section 1. Section 2 describes how each shock is introduced within this framework. Section 3 reports the results in terms of the costs of each shock. Section 4 concludes.

# 1. A LIFE-CYCLE MODEL OF CONSUMPTION AND **NET WORTH**

The economy is that of Athreya (2008), and consists of a continuum of J overlapping generations of working households. Households value consumption, do not value leisure, and therefore supply labor inelastically.

$\begin{array}{c} \mathbf{Parameter} \\ \alpha \end{array}$		Parameter $R^f$		Parameter γ		Parameter $\underline{\tau}$	<b>Value</b> \$7,600
β	0.96	$\psi$	3.4%	$\sigma_u^2$	0.063	$\underline{\tau}^R$	\$8,600
$x_1$	0	λ	0.9	$\sigma_{\eta}^2$	0.0275		
				$\sigma_{\eta 1}^2$	0.22		

**Table 1 Model Parametrization** 

### **Preferences**

The household chooses consumption,  $\{c_j\}_{j=1}^J$ , and retirement wealth,  $x_R$ , to solve

$$\sup_{(\{c_j\}, x_R) \in \Pi(\Psi_0)} E_0 \sum_{j=1}^J \beta^j \frac{c_j^{1-\alpha}}{1-\alpha} + \frac{x_R^{1-\alpha}}{1-\alpha}.$$
 (1)

Here,  $\Pi(\Psi_0)$  denotes the space of all feasible combinations ( $\{c_j\}$ ,  $x_R$ ) given initial state  $\Psi_0$ ,  $\alpha$  denotes risk aversion, and  $\beta$  is the discount factor. In the calibration, risk aversion and the discount factor are set at the standard values of  $\alpha=2$  and  $\beta=0.96$ . (See Table 1 for all model parameter values, which follows Athreya [2008].)

#### **Income**

Households have three potential sources of income: labor income, meanstested transfer income, and retirement income, with labor income being subject to shocks drawn from a probability structure that is known perfectly by the agent.

#### Labor Income

The model period is one calendar year. Households begin working life at age 21 and retire at age 65. Households face uncertainty in their labor income because of stochastic productivity shocks to their labor supply. Following the literature (e.g., Hubbard, Skinner, and Zeldes 1995; Huggett and Ventura 2000; Storesletten, Telmer, and Yaron 2004), the evolution of log income,  $\ln y_j$ , is modeled as

$$ln y_j = \mu_j + z_j + u_j,$$
(2)

where  $\mu_j$  is an age-specific mean of log income,  $z_j$  is the persistent shock, and  $u_j$  is the transitory shock.

The profile  $\{\mu_j\}_{j=1}^J$  is parameterized using data on the median earnings of U.S. males from the 2000 Census.<sup>2</sup>

The persistent shock,  $z_i$ , is given by

$$z_j = \gamma z_{j-1} + \eta_j, \ \gamma \le 1, \ j \ge 2, \eta_j \sim i.i.d. \ N(0, \sigma_n^2).$$
 (3)

We set  $\gamma=0.99$  and  $\sigma_{\eta}^2=0.0275$  to capture the facts that, in the data, the cross-sectional variance in log income increases substantially, and roughly linearly, over the life cycle; it is roughly 0.28 among 21-year-olds and roughly 0.90 among new retirees. The transitory shock,  $u_j$ , is distributed as  $u_j \sim i.i.d\ N(0,\sigma_u^2)$  and is independent of  $\eta_j$ .

To capture initial heterogeneity across households, it is assumed that they draw their first realization of the persistent shock from a distribution with a different variance than at all other ages. That is,

$$z_0 = 0$$
, and  $\eta_1 \sim N(0, \sigma_{\eta_1}^2)$ . (4)

In the above,  $\sigma_{\eta_1}^2 = 0.22$ .

Note that the assumption that households supply labor inelastically restricts them from using a smoothing mechanism that could be particularly useful in the face of unanticipated shocks. However, not only does this assumption keep the model parsimonious, it is in keeping with the usefulness of providing an upper bound on the costs of the shocks we study.

## Means-Tested Transfer Income

Following Hubbard, Skinner, and Zeldes (1995), means-tested transfers  $\tau(\cdot)$  are specified as a function of age, j, net worth,  $x_j$ , and income,  $y_j$ , as follows:

$$\tau(j, x_i, y_i) = \max\{0, \underline{\tau} - (\max(0, x_i) + y_i)\}. \tag{5}$$

Social insurance in the United States aims to provide a floor on consumption and the specification in equation (5) captures this feature. The transfer scheme provides households with a minimum of  $\underline{\tau}$  units of the consumption good at the beginning of the period. In the calibration,  $\underline{\tau} \cong \$7,600$  to match data on the asset accumulation of households in the lower percentiles of the wealth distribution.

#### Retirement Income

Household utility at retirement is evaluated as  $\frac{x_R^{1-\alpha}}{1-\alpha}$ . Retirement wealth,  $x_R$ , is the sum of household personal savings,  $x_{J+1}$ , and a baseline retirement benefit,  $x_{\tau^R}$ :

$$x_R = x_{J+1} + x_{\tau^R}. (6)$$

<sup>&</sup>lt;sup>2</sup> Since income is lognormally distributed, the mean of log income equals the log of median income. Therefore, the log of median earnings is used to generate the profile.

The baseline retirement benefit,  $x_{\underline{\tau}^R}$ , yields an annual income of  $\underline{\tau}^R$  when annuitized using discount rate  $R^f$ . That is, the baseline retirement benefit solves

$$\sum_{k=1}^{K} \frac{\underline{\tau}^{R}}{(R^{f})^{k}} = x_{\underline{\tau}^{R}}.$$
 (7)

Here,  $\underline{\tau}^R$  represents the societal minimum amount of consumption at retirement. This amount is not means tested and is intended to represent the sum of welfare programs, Social Security, and Medicare.<sup>3</sup> The interest rate,  $R^f > 0$ , is the risk-free rate of return on savings and is exogenously given.

In the calibration, the minimum amount of consumption at retirement,  $\underline{\tau}^R$ , is set equal to \$8,600 and  $R^f = 1.01$ .

# **Technology and Market Arrangement**

At each age j, households choose whether to save  $(x_{j+1} > 0)$  or borrow  $(x_{j+1} < 0)$ . Savings earn the exogenous risk-free rate of return  $R^f > 0$ . The interest rate on borrowing is  $R(\cdot)$ , which incorporates credit risk (because households can default on the debt next period) and transaction costs,  $\psi$ , arising from resources used in intermediation. Default is costly and reduces household utility by  $\lambda$  in the period in which debts are repudiated. This cost includes, but is not limited to, the cost of legal representation and court fees. It is meant to capture all costs deemed relevant by households, and will be calibrated to help the model match default-related behavior.

## **Recursive Formulation**

The household's problem is recursive in a state vector that includes age, j, beginning-of-period net worth,  $x_j$ , current-period realization of the persistent shock,  $z_j$ , and current-period realization of transitory income,  $u_j$ .

#### Value Functions

Households that enter a period with debt must decide whether or not to default. The value function when repaying debts is  $W^R(\cdot)$ , which solves

$$W^{R}(j, x_{j}, z_{j}, u_{j}) = \sup_{x_{j+1}} \left\{ \frac{c_{j}^{1-\alpha}}{1-\alpha} + \beta E_{z_{j+1}|z_{j}} V(j+1, x_{j+1}, z_{j+1}, u_{j+1}) \right\},$$
(8)

<sup>&</sup>lt;sup>3</sup> This approach follows Huggett (1996) and Gourinchas and Parker (2002).

subject to

$$c_j + \frac{x_{j+1}}{R(j, x_{j+1}, z_j)} \le y_j + \tau(j, x_j, y_j) + x_j, \tag{9}$$

where  $R(j, x_{j+1}, z_j)$  is the interest rate associated with the level of savings or borrowing,  $x_{j+1}$ , chosen by the household of age j and current realization of the persistent shock  $z_j$ .

The value of defaulting is given by  $W^D(\cdot)$ , which solves

$$W^{D}(j, x_{j}, y_{j}, u_{j}) = \sup_{x_{j+1}} \left\{ \frac{c_{j}^{1-\alpha}}{1-\alpha} - \lambda + \beta E_{z_{j+1}|z_{j}} V(j+1, x_{j+1}, z_{j+1}, u_{j+1}) \right\},$$
(10)

subject to

$$c_j + \frac{x_{j+1}}{R^f} \leq y_j + \tau(j, x_j, y_j), \tag{11}$$

$$x_{i+1} \geq 0. \tag{12}$$

The debt obligation in the right-hand side of (9) does not appear in (11) because the household defaults. The household pays the utility cost,  $\lambda$ , associated with defaulting. In the parametrization,  $\lambda$  is set to 0.9. Along with the other parameters in the model, this targets the Chapter 7 filing rate of 0.5 percent and the mean net worth of Chapter 7 bankruptcy filers of \$16,815.<sup>4</sup>

The household is not allowed to borrow in the period in which it defaults. Thus, net worth chosen in the current period must be non-negative and earns the risk-free rate of interest,  $R^f$ . The household can borrow in all subsequent periods.

The beginning-of-period value function must therefore satisfy

$$V(j, x_j, z_j, u_j) = \max \left[ W^R(j, x_j, z_j, u_j), W^D(j, x_j, z_j, u_j) \right].$$
 (13)

Once borrowing or savings is chosen, the period ends.

# **Loan Pricing**

In the market for loans, creditors are assumed to be competitive and to hold a sufficiently large number of loans of any given size for the law of large numbers to guarantee them a deterministic rate of return on loans of that size. They pay transactions costs,  $\psi$ , in exchange for which they can observe all factors needed to forecast the risk of default one period ahead. In the model, these factors are age, j, the persistent shock,  $z_j$ , and household debt,  $x_j$ . Creditors expect to break even on each loan by pricing contingent on these factors. Let  $\pi^D(j, x_{j+1}, z_j)$  denote the probability of default on a loan of

<sup>&</sup>lt;sup>4</sup>This data is as of 1991 in order to be consistent with the timing of the income and consumption data.

size  $x_{j+1}$ , made to a household of age j, with persistent income shock  $z_j$ . Let  $I(j+1,x_{j+1},z_{j+1},u_{j+1})$  be the indicator function over whether or not a household with debt  $x_{j+1}$  and shocks  $z_{j+1}$  and  $u_{j+1}$  will choose to default. That is,

$$I(j + 1, x_{j+1}, z_{j+1}, u_{j+1}) = 1,$$

if and only if

$$W^{D}(j+1, x_{j+1}, z_{j+1}, u_{j+1}) > W^{R}(j+1, x_{j+1}, z_{j+1}, u_{j+1}).$$

Therefore,  $\pi^D(\cdot)$  is calculated at each age j as follows:

$$\pi^{D}(j, x_{j+1}, z_{j}) = \int \int I(j+1, x_{j+1}, z_{j+1}, u_{j+1}) f(z_{j+1}, u_{j+1}|z_{j}) dz_{j+1} du_{j+1}.$$
(14)

Given  $\pi^D(\cdot)$ , the interest rate function,  $R(j, x_{j+1}, z_j)$ , is determined as follows:

$$R(j, x_{j+1}, z_j) = \frac{R^f + \psi}{(1 - \pi^D(j, x_{j+1}, z_j))}.$$
 (15)

#### 2. UNANTICIPATED SHOCKS AND THEIR SIZES

We now introduce unanticipated shocks to households in the above framework. We capture the effect of the shock on the "representative" household of any given age, as described by the age-specific median value of wealth.

As mentioned earlier, we quantify the effect of such shocks in terms of annual consumption. There are several ways in which we could do this. For ease of interpretation, we express all quantities in terms of *constant* consumption levels. We now describe the two scenarios under study and detail the particular calculations needed to derive the costs in terms of equivalent constant consumption levels under each.

# Case 1: An Unanticipated Drop in Net Worth

The first case analyzes the consequences of an unanticipated drop in net worth. The empirical parallels we have in mind are unexpected decreases in house prices or stock prices. Since wealthier households are likely to have more expensive homes and larger stock portfolios, we assume that the shock is proportional to net worth.

The cost of this shock is calculated as follows. Let  $V(j, \overline{x}_k, z_j, u_j)$  denote the value of arriving in a given period k with wealth  $\overline{x}_k$ . Let  $V(j, \widetilde{x}_k, z_j, u_j)$  denote the value of arriving in a given period k with wealth  $\widetilde{x}_k$ , where, for  $0 < \theta < 1$ ,

$$\widetilde{x}_k = (1 - \theta)\overline{x}_k \text{ if } \overline{x}_k \ge 0$$
 $\widetilde{x}_k = (1 + \theta)\overline{x}_k \text{ if } \overline{x}_k < 0.$ 

The latter is the value associated with the discounted expected utility that a household can obtain from behaving optimally after the occurrence of the shock to net worth. Given this, we define  $\bar{c}$  and  $\tilde{c}$  as the constant values for consumption that, when received over an entire lifetime, generate discounted utility equal to  $V(j, \bar{x}_k, z_j, u_j)$  and  $V(j, \tilde{x}_k, z_j, u_j)$ , respectively. Thus,  $\bar{c}$  and  $\tilde{c}$  solve, respectively,

$$V(j, \overline{x}_k, z_j, u_j) = \sum_{k=j}^{J+25} \beta^{j-k} \frac{\bar{c}^{1-\alpha}}{1-\alpha},$$

$$V(j, \widetilde{x}_k, z_j, u_j) = \sum_{k=j}^{J+25} \beta^{j-k} \frac{\tilde{c}^{1-\alpha}}{1-\alpha}.$$

The difference  $\bar{c} - \tilde{c}$  represents the number of units of consumption that would make the household indifferent to facing the shock or not. The difference thus represents the cost of the shock in units of consumption, or, alternately, the amount the household would pay in units of consumption to avoid facing the shock. (See the Appendix for calculation details.)

What are empirically plausible sizes of the net worth shocks? We may arrive at upper bounds by assuming that the household's entire net worth is composed of a single asset and attribute the shock to a drop in the value of that asset. This asset might be the household's equity in its home or its stock portfolio. To obtain an approximation of the upper bound on the size of shocks to house and stock prices, we use available data. The recent house price bust serves as a case in point. The largest annual decline of nearly 19 percent in the S&P/Case-Shiller Home Price Index of U.S. National Values since 1987 came between the first quarter of 2008 and the corresponding quarter of 2009.<sup>5</sup> If we look at stocks, the shock could correspond to the worst one-year performance of diversified mutual funds such as Vanguard's S&P 500 Index or Total Stock Market Index, in which their value fell by roughly 37 percent. Different values may be appropriate for households with different profiles, but we carry out the exercise for a 40 percent drop in asset values, i.e.,  $\theta = 0.4$ , to serve as an upper bound.

## Case 2: An Unexpected Tightening of Credit Markets

Our second scenario aims to measure the cost imposed on a household by a sudden change in borrowing premia, and so intends to be reflective of dysfunction in credit markets very generally. We capture this in the model by comparing the maximal value that is attainable to an agent under the initial

<sup>&</sup>lt;sup>5</sup> See

www.standardandpoors.com/indices/sp-case-shiller-home-price-indices/en/us/?indexId=spusa-cashpidff-p-us---.

interest rate function with that attainable from living in an environment where loans are costlier than before. Specifically, we model the shock as raising the interest rate on credit to  $R(\cdot) + \iota$ ,  $\iota > 0$ . Importantly, we assume that agents will be faced with the tighter credit conditions *for the rest of their lives*. As a result, our calculations will likely represent an upper bound on the cost of such credit market tightening.

The net worth shock did not change household value functions. This is because the shock did not alter the subsequent uncertainty or costs in the household's environment. In this case, by changing the interest rate faced by households for the rest of their lives, the shock does change the maximal value of attainable utility coming from any given wealth position. Let  $V(j, x_k, z_j, u_j)$  be the value function before the shock, when the interest on credit is  $R(\cdot)$ . Let  $\widetilde{V}(j, x_k, z_j, u_j)$  be the value function after the shock, derived from solving the household's optimization problem over its remaining life under the new credit-pricing function.

In this case,  $\bar{c}$  and  $\tilde{c}$  solve

$$V(j, x_k, z_j, u_j) = \sum_{k=j}^{J+25} \beta^{j-k} \frac{\bar{c}^{1-\alpha}}{1-\alpha},$$

$$\sum_{k=j}^{J+25} \beta^{j-k} \tilde{c}^{1-\alpha}$$

$$\widetilde{V}(j, x_k, z_j, u_j) = \sum_{k=j}^{J+25} \beta^{j-k} \frac{\widetilde{c}^{1-\alpha}}{1-\alpha}.$$

As before, cost of shock is  $\bar{c} - \tilde{c}$ .

We increase borrowing costs at all debt levels by 300 basis points, which corresponds to among the largest spreads observed between mortgage market interest rates and 10-year Treasury securities.

## 3. RESULTS

The following section presents results for the two cases. In both scenarios, we impose the shock on households at one period and calculate the cost of the shock for households of various ages. The household does not expect, nor does it receive, any further shocks aside from the age at which we study them.

Notice that in our model, the presence of uninsurable risk will lead households to vary not only in income, but importantly, in consumption and wealth. There is, therefore, no "representative" household. This raises the issue of *whose* well-being and costs we are studying. A natural candidate, which we use, is the household with the median level of wealth and the median level of income shocks for its age group. Figure 1 shows median wealth for households at each age.

We find in general that the cost depends on the household's initial level of assets, the size of the shock, and the time period in which it occurred. Each case is discussed in detail.

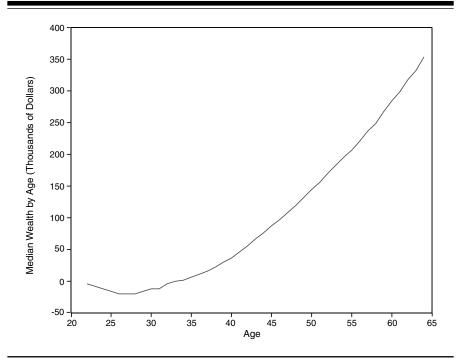


Figure 1 Median Wealth by Age

#### Case 1: Net Worth Shock

We calculate the cost of a 40 percent decrease in net worth to a household with the median level of assets and income shocks for its age. Figure 2 shows the cost of the shock to the household in dollars of constant annual consumption, which is calculated as  $\tilde{c} - \bar{c}$ . Figure 3 shows the cost as a fraction of constant annual consumption, calculated as  $\frac{\tilde{c}-\bar{c}}{\bar{c}}$ .

The cost as a function of age displays a U-shape and then a steep increase. The U-shape corresponds to ages at which the household has negative net worth. The proportional shock pushes the household deeper into debt, and is costliest when the household has the most debt (at age 26). Subsequently, the cost of the shock is small for a while. This is because the absolute value of wealth is low, so a proportional decline amounts to a very small number. However, the cost rapidly increases with age. As age increases, so does median wealth and the same percentage drop in net worth represents a much larger absolute loss. For example, the cost to a household that faces the shock at age 40 is \$481 in constant annual consumption for the rest of its life.<sup>6</sup> For

<sup>&</sup>lt;sup>6</sup> All costs in this article are reported in 2010 dollars.

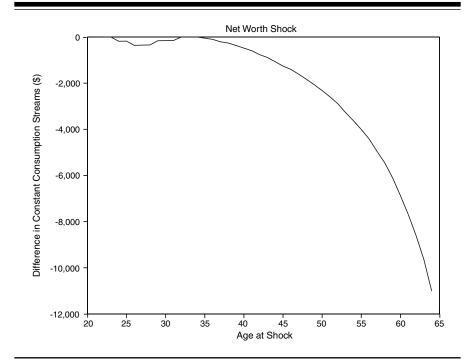


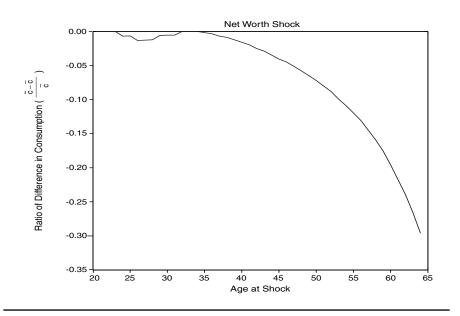
Figure 2 Difference in Consumption by Age at Shock: Case 1

a household at age 60, the cost of an unforeseen 40 percent reduction in net worth is \$6,870 per year for the rest of its life. Figure 3 represents the same cost in percentage terms. The cost to the 40-year-old household is 1 percent of annual consumption without the shock, while for the 60-year-old household, it is nearly 15 percent.

While older households face greater costs in terms of annual consumption, they also face them for fewer periods. It is therefore useful to compare the present value of the sequence  $\{\bar{c} - \tilde{c}\}_{j=k}^{J+25}$  for various k, where k is the date of the shock. Calculating the present value in the first period of the model gives the perspective of one household looking ahead, while calculating the present value at date k compares the relative cost of the shock to older and younger households living in date k. Figure 4 shows the results of both calculations. Even in present value terms, the cost of the shock is highest for the oldest households.

<sup>&</sup>lt;sup>7</sup> See the Appendix for calculation details.

Figure 3 Percentage Difference in Consumption by Age at Shock: Case 1



#### Case 2: Interest Rate Shock

We now consider the shock to a household that comes from facing unexpectedly higher interest rates on credit for the rest of its working life. Recall that in this case, we measure the cost as arising from a 300-basis-point increase in the interest rate on all debt levels that the household might choose. The corresponding costs in terms of annual consumption, fraction of original consumption, and present value of annual consumption are shown in Figures 5–7.

The cost of this shock is very small relative to the net worth shock; it does not exceed \$300 of annual consumption for agents of any age. An important part of why this cost remains small is that households can adjust savings in the interim fairly effectively to nullify the effects of such an increase in costs. Moreover, as long as such a shock does not occur at the time when a household is holding peak debt (lowest net worth), the size of the shock itself is not large. Lastly, once households leave the first 15 years or so of their working life cycle, they are typically not in debt (have non-negative net worth), and, moreover, do not typically expect to return to such a state. Thus, contractions in credit markets will not hurt them.

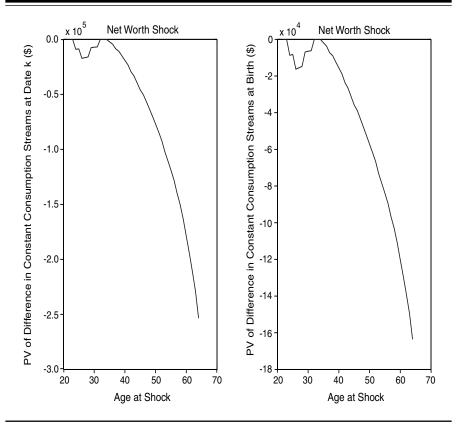


Figure 4 Present Value of Difference in Consumption: Case 1

The costs we report have all been calculated under the assumption that households can declare bankruptcy and remove all unsecured debt obligations subject to a penalty. As a result, debt in the model is priced to reflect this possibility. Household net worth over the life cycle is, of course, different than what it would be in a setting where households did not have this option but instead had access to risk-free borrowing. As a check for robustness, we have shut down this option in the model by making the utility cost,  $\lambda$ , infinite, which effectively precludes bankruptcy. We find that this has little or no effect on the size of the costs. This is because the option to declare bankruptcy is relevant only to a subset of households—those with negative net worth (that do not have sufficient assets to pay off their debts). In our model, these are younger households. Because the debt they hold is not large on average, the proportional net worth shock translates into a small shock for them in absolute

<sup>&</sup>lt;sup>8</sup> Results are available from the authors upon request.

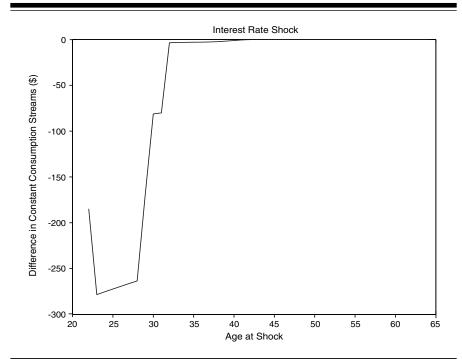


Figure 5 Difference in Consumption by Age at Shock: Case 2

terms. Our results have shown that the cost of the interest rate shock is also small in absolute terms for these households. As a result, neither shock is large enough to make bankruptcy an important consideration for the examples we study.

We note here that the focus of the model on net worth is likely very important for the small role it assigns to the effects of an interest rate shock. In a richer setting, the fact that households are often engaged in very heavily leveraged investment (taking out mortgages to finance a home purchase) means that credit market costs could likely affect people well into late middle-age. This is simply because, while they might have positive net worth by middleage (indeed will, in most instances), they may also owe substantial amounts on a mortgage, and the size of these obligations may be quite large and difficult to deal with. In ongoing work (Athreya, Ionescu, and Neelakantan 2011), we are considering precisely this.

#### **CONCLUSION**

In this article, we take a first step in measuring the cost of two particular and, we think, representative types of financial shocks. The results yield some general

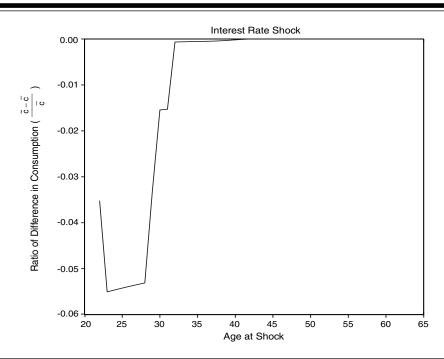


Figure 6 Percentage Difference in Consumption by Age at Shock: Case 2

insights about such shocks and their costs. Comparing across households of various ages, shocks that are proportional to net worth are costliest to the oldest households for which the proportional shock translates into the largest absolute drop in net worth. Interest rate shocks, in the form of an unanticipated tightening of the credit market, are much less costly.

As mentioned at the outset, this article is stylized along several dimensions, and thus represents only a small first step in the important task of assessing the power of financial shocks to compromise household well-being. In particular, the model abstracts from portfolio choice and focuses instead on a simple scalar measure of net worth. This in turn prevents us from fully analyzing particular kinds of financial decisions, such as house purchases or any other leveraged purchase of risky assets, which can greatly change gross financial positions and periodic payment obligations while leaving net worth essentially unchanged. The model also abstracts from the labor supply decision, which could mitigate the cost of the shocks. Finally, the shocks that the household faces are completely unanticipated, something that is likely not as stark in reality. Households may well be aware of the existence of the kinds of shocks we analyze in this article, but incorrect about or unable to assess

Interest Rate Shock Interest Rate Shock PV of Difference in Constant Consumption Streams at Date k (\$) PV of Difference in Constant Consumption Streams at Birth (\$) -2,000 -2,000 -4,000 -4,000 -6,000 -6,000 -8,000 -8,000 -10,000 -10,000 -12,000 -12,000 -14,000 -14,000 30 40 50 60 30 40 50 60 70 Age at Shock Age at Shock

Figure 7 Present Value of Difference in Consumption: Case 2

"true" probabilities for such events. In ongoing work (Athreya, Ionescu, and Neelakantan 2011), we consider these issues in a richer model of household portfolio choice.

# **APPENDIX**

Solving for  $\bar{c}$  and  $\tilde{c}$ 

$$V(j, \overline{x}_k, z_j, u_j) = \sum_{j=k}^{J+25} \beta^{j-k} \frac{\overline{c}^{1-\alpha}}{1-\alpha}.$$

First, we change the index of summation. Let i = j - k. Then

$$V(j, \overline{x}_k, z_j, u_j) = \sum_{i=0}^{J+25-k} \beta^i \frac{\overline{c}^{1-\alpha}}{1-\alpha}.$$

We use the following rule for the sum of a finite series:

$$\sum_{i=0}^{n} a^{i} = \frac{1 - a^{n+1}}{1 - a},$$

to obtain

$$V(j, \overline{x}_k, z_j, u_j) = \frac{\overline{c}^{1-\alpha}}{1-\alpha} \left[ \frac{1 - \beta^{J+25-k+1}}{1-\beta} \right].$$

Let

$$\Phi_k = \left\lceil \frac{1 - \beta^{J+25-k+1}}{1 - \beta} \right\rceil.$$

Then

$$V(j, \overline{x}_k, z_j, u_j) = \frac{\overline{c}^{1-\alpha}}{1-\alpha} \Phi_k.$$

Solving for  $\bar{c}$  yields

$$\bar{c} = \left[ \frac{V(j, \overline{x}_k, z_j, u_j)}{\Phi_k} (1 - \alpha) \right]^{\frac{1}{1 - \alpha}}.$$
 (16)

Similarly,

$$\tilde{c} = \left[ \frac{V(j, \widetilde{x}_k, z_j, u_j)}{\Phi_k} (1 - \alpha) \right]^{\frac{1}{1 - \alpha}}.$$
(17)

# Finding the Present Value of $\bar{c} - \tilde{c}$

To allow us to easily compare the cost of shocks at various ages, we now compute two types of present values. First, we begin by discounting to age  $\theta$ , not just back the date at which the shock occurred (date k). The present value at date zero of a shock occurring at date k, given that the constant consumption equivalents are  $\bar{c}$  and  $\tilde{c}$ , is

$$PV_0(k) = \frac{\bar{c} - \tilde{c}}{(1+r)^k} + \frac{\bar{c} - \tilde{c}}{(1+r)^{k+1}} + \dots + \frac{\bar{c} - \tilde{c}}{(1+r)^{J+25}},$$

where r is the interest rate on savings. To be clear, notice that the first discounting term  $\frac{1}{(1+r)^k}$  shows that we are discounting to age 0 events that begin at age k.

Let  $\frac{1}{1+r} = \hat{\beta}$ . Next, we'll use the known formula for the finite sum of a geometric series. We want the sum from age k to death (age J+25). We

therefore first take the series from 0 to J+25,  $\left[\frac{1-\hat{\beta}^{J+25+1}}{1-\hat{\beta}}\right]$  and subtract from this the sum going from 0 to k-1,  $\left[\frac{1-\hat{\beta}^k}{1-\hat{\beta}}\right]$ .

$$\begin{split} PV_0(k) &= (\bar{c} - \tilde{c}) \left[ \frac{1 - \hat{\beta}^{J+25+1}}{1 - \hat{\beta}} - \frac{1 - \hat{\beta}^k}{1 - \hat{\beta}} \right] \\ &= (\bar{c} - \tilde{c}) \left[ \frac{\hat{\beta}^k - \hat{\beta}^{J+25+1}}{1 - \hat{\beta}} \right] \\ &= (\bar{c} - \tilde{c}) \left[ \frac{1 - \hat{\beta}^{J+25-k+1}}{1 - \hat{\beta}} \right] \hat{\beta}^k \\ &= (\bar{c} - \tilde{c}) \Phi_{diff\_at\_birth}, \end{split}$$

letting

$$\Phi_{diff\_at\_birth} = \left[\frac{1 - \hat{\beta}^{J+25-k+1}}{1 - \hat{\beta}}\right] \hat{\beta}^k,$$

we have

$$PV_0(k) = (\bar{c} - \tilde{c})\Phi_{diff\_at\_birth}.$$

Similarly, the present value of a shock occurring at date k, discounted back to date k is

$$PV_k(k) = (\bar{c} - \tilde{c})\Phi_{diff\_at\_k},$$

where we define

$$\Phi_{diff\_at\_k} = \left[ \frac{1 - \hat{\beta}^{J+25-k+1}}{1 - \hat{\beta}} \right].$$

## REFERENCES

Athreya, Kartik B. 2008. "Default, Insurance, and Debt over the Life-Cycle." *Journal of Monetary Economics* 55 (May): 752–74.

Athreya, Kartik B., Felicia A. Ionescu, and Urvi Neelakantan. 2011. "How Nasty are Financial Surprises? A View Over the Life-Cycle." Mimeo, Federal Reserve Bank of Richmond.

Baxter, Marianne, and Robert G. King. 1993. "Fiscal Policy in General Equilibrium." *American Economic Review* 83 (June): 315–34.

- Gourinchas, Pierre-Olivier, and Jonathan A. Parker. 2002. "Consumption Over the Life Cycle." *Econometrica* 70 (January): 47–89.
- Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes. 1995. "Precautionary Saving and Social Insurance." *Journal of Political Economy* 103 (April): 360–99.
- Huggett, Mark. 1996. "Wealth Distribution in Life-Cycle Economies." *Journal of Monetary Economics* 38 (December): 469–94.
- Huggett, Mark, and Gustavo Ventura. 2000. "Understanding Why High Income Households Save More than Low Income Households." *Journal of Monetary Economics* 45 (April): 361–97.
- Joffe-Walt, Chana, and Adam Davidson. 2010. "Housing Guy Apologizes for Housing Bubble." Available at www.npr.org/blogs/money/2010/10/22/ 130756958/finance-guy-apologizes-for-housing-bubble.
- Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron. 2004. "Consumption and Risk Sharing Over the Life Cycle." *Journal of Monetary Economics* 51 (April): 609–33.