

Generalized Matching Functions and Resource Utilization Indices for the Labor Market

Andreas Hornstein and Marianna Kudlyak

In the years following the Great Recession, the signals for a recovery of the U.S. labor markets were mixed: while the unemployment rate declined to historically low levels, labor force participation rates also declined. This observation raised doubts on the ability of the unemployment rate alone to accurately represent the state of resource utilization in the labor market.¹ In Hornstein, Kudlyak, and Lange (2014), we therefore proposed an indicator of resource utilization in the labor market, a nonemployment index (NEI), that is more comprehensive than the standard unemployment rate. In this article, we relate our NEI to recent research on frictional unemployment in labor markets and thereby provide a theoretical grounding for the NEI beyond the heuristic justifications for its usefulness in our previous work.

More than 30 years ago, Flinn and Heckman (1983) pointed out that the distinction between those being unemployed and those being out of the labor force (OLF) is not clear cut but a matter of degree. For example, the unemployed, that is, those nonemployed who are actively searching for work, are twice as likely to make the transition to

■ Andreas Hornstein is a senior advisor at the Federal Reserve Bank of Richmond and Marianna Kudlyak is a senior economist in the Research Department at the Federal Reserve Bank of San Francisco. The authors thank Sean McCrary for excellent research assistance and Marios Karabarbounis, Santiago Pinto, Allen Sirolly, and John Weinberg for helpful comments. The views expressed in this article are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond, the Federal Reserve Bank of San Francisco, or the Federal Reserve System. E-mail: Andreas.Hornstein@rich.frb.org; Marianna.Kudlyak@sf.frb.org.

¹ See, for example, Appelbaum (2014), Yellen (2014), or Irwin (2017).

employment within a month than those nonemployed who express a desire to work but do not actively engage in job search activities, and they are three times as likely to make the transition to employment than those who do not even express a desire to work. Thus even though the differences in employment transition probabilities are quantitatively large, they do not suggest a qualitative difference between being unemployed and being OLF. Furthermore, despite the substantially lower employment transition probabilities for OLF, on average, every month twice as many people make the transition from OLF to employment than do from unemployment.

The Diamond-Mortensen-Pissarides search-matching framework interprets new employment as being “produced” by matching job seekers with open positions.² The standard approach assumes a homogeneous search pool, that is, each searcher is equally likely to make the transition to employment. Recent extensions have emphasized the heterogeneous nature of the search pool, that is, the persistent differences in search efficiency between unemployment and OLF, which is reflected in persistent differences of employment transition probabilities, for example, in Veracierto (2011), Diamond (2013), Elsby, Hobijn, and Şahin (2015), Barnichon and Figura (2015), and Hornstein and Kudlyak (2016). Most of this work is done in the context of estimating matching efficiency in the labor market, that is, the extent of labor market frictions. Accounting for heterogeneity in the search pool leads to smaller estimated declines in matching efficiency, in part since heterogeneity introduces systematic positive comovement between total nonemployment and the average search efficiency of the heterogeneous search pool. Within this generalized matching framework, we interpret our proposed NEI as the quality-adjusted measure of the search pool.

This article is structured as follows. We first review the search-matching framework and how it accounts for changes in average employment transition rates with homogeneous and heterogeneous search pools. We then characterize the pool of nonemployed in the Current Population Survey (CPS) in terms of their average transition rates to employment. Finally, we construct a sequence of NEIs with increasing coverage of the nonemployed, the most comprehensive of them being the NEI proposed in Hornstein et al. (2014). We show how these NEIs fit into a generalized search-matching framework with heterogeneous search pools and study their implications for measured changes in matching efficiency. We should note that there is substantial overlap

² For example, Pissarides (2000) or Petrongolo and Pissarides (2001).

between this paper and Hornstein et al. (2014), especially as it relates to the characterization of the nonemployed in the CPS.

1. GENERALIZED MATCHING FUNCTIONS

The aggregate search and matching function in macro-labor models describes the “production” of hires as a function of the stocks of job seekers and vacancies and an exogenous shift term denoting the aggregate efficiency of the matching process. The standard approach for the search and matching function assumes that the inputs are homogeneous. We augment the standard search and matching function by allowing for fixed heterogeneity across observed groups of job seekers.

The Matching Function with Homogeneous Search

Consider an economy where unemployed workers need to be matched with open positions. Assume that all workers and open positions are homogeneous, but that for some reason the assignment of unemployed workers to open positions is a time-consuming process. This process is characterized by a matching function,

$$h = e^{\kappa} v^{\alpha} u^{1-\alpha}, \quad (1)$$

where h is the number of new hires when v vacancies are matched with u unemployed workers, and $\alpha \in [0, 1]$ is the elasticity of new hires with respect to vacancies. The matching function is constant returns to scale, that is, if the number of vacancies and unemployed doubles, then the number of new matches also doubles. In fact, the usual specification of the matching function in equation (1) is analogous to a Cobb-Douglas production function where unemployed workers and vacancies are inputs to a process that generates new filled positions. This process may be more or less efficient, and the matching efficiency κ reflects the extent of frictions in the labor market. The smaller the matching efficiency, the less efficient the labor market is at matching the unemployed with open positions.

The rate at which unemployed workers make the transition to employment is

$$\lambda = \frac{h}{u} = e^{\kappa} \left(\frac{v}{u} \right)^{\alpha} = e^{\kappa} \theta^{\alpha}, \quad (2)$$

where the vacancy-unemployment ratio θ denotes “labor market tightness.”³ Conditional on the matching elasticity, we can recover the matching efficiency from observations on how long it takes for an unemployed to become employed, that is, the employment transition rate and market tightness,

$$\kappa = \ln \lambda - \alpha \ln \theta. \quad (3)$$

Heterogeneous Search

Now suppose that the unemployed differ in their search effectiveness, but that after accounting for these differences, they are all perfect substitutes in the matching function. First assume that there is a finite number of types, J , and that each type is endowed with ρ_j search units. The total effective search input from all of these types is

$$u^* \equiv \sum_{j=1}^J \rho_j u_j, \quad (4)$$

and together with the available vacancies the matching function determines total hired search units

$$h = e^\kappa v^\alpha (u^*)^{1-\alpha}.$$

Analogous to the case of homogeneous searchers, the rate at which a search unit will make the transition to employment is then

$$\lambda^* = e^\kappa \left(\frac{v}{u^*} \right)^\alpha.$$

Since a type j agent is endowed with ρ_j search units, her employment transition rate is

$$\lambda_j = \rho_j \lambda^*,$$

and the differences in search effectiveness account for differences in employment transition rates across types.

We can relate this simple model of search heterogeneity to the baseline model with homogeneous search by explicitly accounting for the average search effectiveness across types,

$$\bar{\rho} = \sum_j \frac{u_j}{u} \rho_j. \quad (5)$$

³ We interpret the transitions as occurring continuously over time. In particular, we assume that employment opportunities arrive according to a Poisson process with arrival rate λ . In this case, a worker who is unemployed at the beginning of the period will be employed at the end of the period with probability $1 - e^{-\lambda}$. See also the Appendix.

The employment transition rate per search unit is then

$$\lambda^* = e^\kappa (\theta \bar{\rho})^\alpha,$$

and the average employment transition rate across all types is

$$\bar{\lambda} = \sum_j \frac{u_j}{u} \rho_j \lambda^* = e^\kappa \theta^\alpha \bar{\rho}^{1-\alpha}. \quad (6)$$

Thus, we have to correct for changes in average search effectiveness when we recover the matching efficiency from observations on the average employment transition rate and market tightness,

$$\kappa = \ln \bar{\lambda} - \alpha \ln \theta - (1 - \alpha) \ln \bar{\rho}. \quad (7)$$

In other words, assuming that all workers in the search pool are homogeneous when they are not conflates changes in matching efficiency with changes in average search effectiveness.

2. HETEROGENEITY OF NONEMPLOYMENT

We now briefly describe the components of nonemployment that we use in the construction of our nonemployment index. This section is closely related to Section 1 of Hornstein, Kudlyak, and Lange (2014).

The BLS Classification Scheme

Among the most widely reported statistics from the Bureau of Labor Statistics (BLS) are the shares of the working-age population who are currently employed, unemployed, and OLF. These shares are estimated using responses from the monthly CPS. A nonemployed respondent is counted as unemployed if she has been actively looking for work in the month preceding the survey week. Those neither employed nor actively looking for work are classified as OLF. Starting with the comprehensive revision of the CPS in 1994, the BLS provides additional detail on the labor market attachment of the nonemployed based on survey responses as to why an individual is not actively looking for work (see Polivka and Miller [1998] for a description of the 1994 CPS revision). The average population shares for the different nonemployment categories in the CPS are listed in Table 1, in columns 1a and 1b. We report the average shares for the years 1994–2007 in column 1a and for the years 2008–16 in column 1b. The first sample represents a relatively strong labor market: it includes two expansions, in particular, the late 1990s information technology boom, and the shallow 2001 recession. The second sample is dominated by the 2008–09 Great Recession and represents a relatively weak labor market.

Table 1 Nonemployment by BLS Categories

	WAP Share		Transition Probability			
	(1a)	(1b)	(2a) ^{PE}	(2b)	(3a) ^{PNE}	(3b)
Unemployed						
Short Term	2.8	3.0	30.2	24.3	26.8	28.2
	[7.6]	[7.5]	[1.00]	[1.00]	[1.00]	[1.00]
Long Term	0.6	1.6	16.0	11.3	29.6	25.7
	[1.6]	[4.0]	[0.53]	[0.47]	[1.10]	[0.91]
OLF, Want to Work						
Marginally attached,	0.2	0.3	13.9	11.8	75.4	74.7
discouraged	[0.5]	[0.8]	[0.46]	[0.48]	[2.81]	[2.64]
Marginally attached,	0.4	0.4	13.7	11.0	73.8	76.5
other	[1.0]	[1.0]	[0.45]	[0.45]	[2.75]	[2.71]
Other	1.7	1.7	15.4	12.8	62.1	65.5
	[4.8]	[4.2]	[0.51]	[0.53]	[2.32]	[2.32]
OLF, Do Not Want to Work						
In school, aged 16-24	3.8	4.9	9.4	6.6	15.2	12.5
	[10.3]	[12.1]	[0.31]	[0.27]	[0.57]	[0.44]
Not in school,	7.6	7.0	7.7	7.2	18.2	20.3
disabled or retired	[20.7]	[17.4]	[0.26]	[0.30]	[0.68]	[0.72]
Disabled	4.2	5.3	1.8	1.5	3.9	4.9
	[11.6]	[13.1]	[0.06]	[0.06]	[0.14]	[0.18]
Retired	15.4	16.2	1.4	1.4	2.0	2.2
	[42.1]	[39.9]	[0.05]	[0.06]	[0.08]	[0.08]
Total	36.6	40.5	6.8	5.8	13.2	13.7
Average						

Note: For the different nonemployed population groups columns 1 display their average percentage shares in total working-age population (WAP). For columns 1, the terms in square brackets represent the nonemployed groups' percentage shares in total nonemployment. Columns 2 display the groups' average transition probabilities to employment, and columns 3 display their average transition probabilities to any other nonemployment state. For the employment transition probabilities, the terms in square brackets represent the average of transition probabilities when normalized with the transition probability of short-term unemployed. Columns a cover the time period 1994–2007 and columns b the time period 2008–16.

The unemployed can be subdivided based on their reported length of unemployment. Short-term unemployment (STU) covers those who have been unemployed for 26 or fewer weeks, while long-term unemployment (LTU) encompasses those who have been unemployed for more than 26 weeks. Prior to the Great Recession, on average, less than one-fifth of all unemployed report more than 26 weeks of unemployment in any one month. But the unemployed represent only one-tenth of the nonemployed. The remaining nine-tenths are OLF.

A little less than one-tenth of the OLF declare that they do want to work, even though they did not actively look for work in the

previous month. Those in this group who want a job, are available for work, and searched for work within the last year (not the last month) are classified as marginally attached. On average, about one-fourth of those who want work are marginally attached, and there are six times as many unemployed as there are marginally attached respondents. Those marginally attached who did not search for a job during the last month because they were discouraged over job prospects are classified as discouraged. On average, discouraged individuals make up about one-third of the marginally attached. But over nine-tenths of those OLF do not want a job. Among these individuals we can distinguish between those who are retired, disabled, currently in school, and the remainder. On average, the retired and disabled account for about two-thirds of those who do not want work.

Despite the recent decline of unemployment to historically low levels in 2016, in the aftermath of the 2007–09 recession average nonemployment is about 4 percentage points higher than it was prior to the recession. Comparing columns (1a) and (1b) of Table 1, we see that the main drivers of this increase of nonemployment were higher LTU, disability and retirement, and more people in school, whereas the share of those OLF who want to work remained relatively stable. The share of LTU increased to close to one-half of total unemployment and has remained high even though overall unemployment has declined. Some of the increase in disability may be in response to the weak labor market of the Great Recession, but it also reflects the continuation of a positive trend established in prior years. Finally, the increased retirement share reflects the demographics of an aging U.S. population.

Transitions to Employment

We are motivated to examine broader nonemployment concepts since the distinction between unemployment and OLF is not as sharp as one would think. In fact, from month to month, roughly twice as many individuals transition from OLF to employment as transition from unemployment. We now show that for all of our nonemployment groups, the transition probabilities to employment are positive and that the heterogeneity in these transition probabilities seems to be consistent with the self-reported labor market attachment.

We first use the CPS microdata to construct exit probabilities from nonemployment using the short rotating four-month panels in the CPS. In any month, we observe the labor market status in the current and following month for roughly three-fourths of the sample. Based on the responses to the CPS questions, we group the nonemployed into the nine nonemployment segments discussed above: the two duration

segments of the unemployed, the three segments of OLF who want a job (marginally attached, discouraged, other), and the four segments of OLF who do not want a job (retired, disabled, in school, not in school). We then construct the transition probabilities into employment or a different nonemployment state for each segment by matching the individual records from the CPS microdata month to month.⁴ The transition probability from a particular segment of nonemployment is the fraction of that segment that exits to employment, p_E , or to a different segment of nonemployment, p_{NE} , from one month to the next.

Table 1, column 2, shows annual averages of the monthly employment transition probabilities for the two unemployment segments and seven OLF segments averaged across 1994–2007 and 2008–16. The chances of becoming employed differ substantially among these groups. The employment probabilities are highest for the short-term unemployed: on average, they have a 30 percent chance of finding a job within a month. Next are the LTU and those OLF individuals who want a job: they are about half as likely to become employed as are the STU.⁵ Then there is the group of those who do not want a job but who are neither retired nor disabled: they are only one-fourth as likely to become employed as are the STU. Finally, there is the group of retired and disabled who are less than one-tenth as likely to become employed as are the STU.⁶

In recessions the employment probabilities tend to fall for all groups, but the ranking of the different groups in terms of their transition probabilities to employment remains the same.⁷ This is also apparent when comparing the pre- and post-Great Recession period, columns 2a and 2b: even though the average transition probabilities are uniformly lower in the post-2008 period, the relative transition probabilities are not that different. Furthermore, the ranking of employment probabilities coincides with the desire to work as stated in the survey: those who actively search tend to have higher transition rates to employment than those who want to work but do not actively look for work, and those who want to work have higher transition rates than those who do not want to work.

⁴ Our matching procedure follows the algorithms described in Madrian and Lefgren (1999) and Shimer (2012). The CPS microdata files are available at http://thedataweb.rm.census.gov/ftp/cps_ftp.html#cpsbasic.

⁵ Note that the employment transition probabilities among the marginally attached OLF do not differ much. In particular, there is no reason to single out discouraged workers based on the likelihood of becoming employed again.

⁶ See also Fujita (2014).

⁷ See Kudlyak and Lange (2014) for graphs of annual averages of monthly job finding rates for the years 1994 to 2013. See also Figures 2 and 3.

Table 1, column 3, shows annual averages of the monthly transition probabilities to a different nonemployment state for the two unemployment segments and seven OLF segments averaged across 1994–2007 and 2008–16. Again, the chances of making the transition to a different nonemployment state differ substantially among these groups, and again the STU stand out. For the STU, the probability of making the transition to a different nonemployment state is slightly lower than the probability of becoming employed, whereas the opposite is true for all other nonemployment states. This is especially noteworthy for those OLF who want to work but are classified as OLF because they do not state that they are actively looking for work. For this group, the probability of exiting to a different nonemployment state is four to five times higher than the probability of becoming employed. It is quite possible that these high probabilities of switching to a different nonemployment state simply mean that individuals in these groups will in the next month state that they are actively looking for work. That being the case, the fact that for all groups except the STU the transition probabilities to some other nonemployment state are higher than the transition probability to employment suggests that looking at the employment transition probability alone as a measure of labor market attachment might be misleading.

We elaborate on the issue of how transition probabilities to employment and some other nonemployment state jointly reflect the transitions to employment in the Appendix. When transitions between employment and nonemployment states take place continuously, the month-to-month transition probabilities that we calculate from the CPS between two points in time reflect this underlying process. In particular, a relatively high transition rate to nonemployment states may mask the true transitions to employment in the employment transition probability. Effectively, the employment transition probability from month to month may appear to be low not because the transition rate to employment is low, but because the transition rate to other nonemployment states with low exit rates to employment is high. In Table 2, we report the employment transition rates using either employment transition probabilities alone in column 1 or transition probabilities to employment and nonemployment jointly in column 2.⁸ Accounting for the interaction between transitions to employment and other nonemployment states tends to increase the estimated level of employment transition rates, but for all nonemployment segments except for the

⁸ In the Appendix, we describe how the transition probabilities can be used to recover the transition rates that generate the observed transition probabilities.

Table 2 Employment Transition Rates by BLS Categories

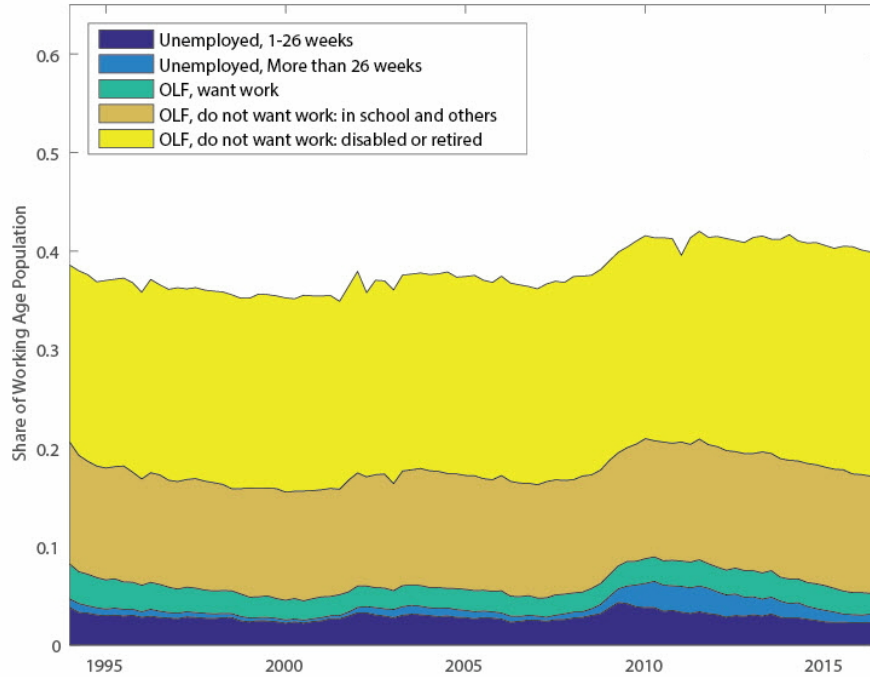
	Employment Transition Rate			
	using p_E (1a)	using p_E (1b)	using p_E and p_{NE} (2a)	using p_E and p_{NE} (2b)
Unemployed				
Short Term	0.36 [1.00]	0.28 [1.00]	0.45 [1.00]	0.35 [1.00]
Long Term	0.17 [0.48]	0.12 [0.43]	0.21 [0.48]	0.14 [0.41]
OLF, Want to Work				
Marginally attached, discouraged	0.15 [0.42]	0.13 [0.45]	0.35 [0.79]	0.27 [0.79]
Marginally attached, other	0.15 [0.41]	0.12 [0.42]	0.33 [0.73]	0.26 [0.76]
Other	0.17 [0.47]	0.14 [0.50]	0.30 [0.67]	0.25 [0.73]
OLF, Do Not Want to Work				
In school, aged 16-24	0.10 [0.28]	0.07 [0.25]	0.11 [0.24]	0.07 [0.21]
Not in school, disabled or retired	0.08 [0.22]	0.08 [0.27]	0.09 [0.20]	0.08 [0.25]
Disabled	0.02 [0.05]	0.02 [0.06]	0.02 [0.04]	0.02 [0.05]
Retired	0.01 [0.04]	0.01 [0.05]	0.01 [0.03]	0.01 [0.04]
Average				
	0.07	0.06	0.08	0.07

Note: For the different nonemployed population groups, columns 1 display the groups' average employment transition rates calculated from employment transition probabilities only, and columns 2 display their average employment transition rates calculated from transition probabilities to employment and any other non-employment state. The details of the employment transition rate calculations are described in the Appendix. The terms in square brackets represent the average of transition rates when normalized with the transition rate of the STU. Columns a cover the time period 1994-2007 and b the time period 2008-16.

OLF who want to work it does not affect the employment transition rates relative to the transition rates of the STU.

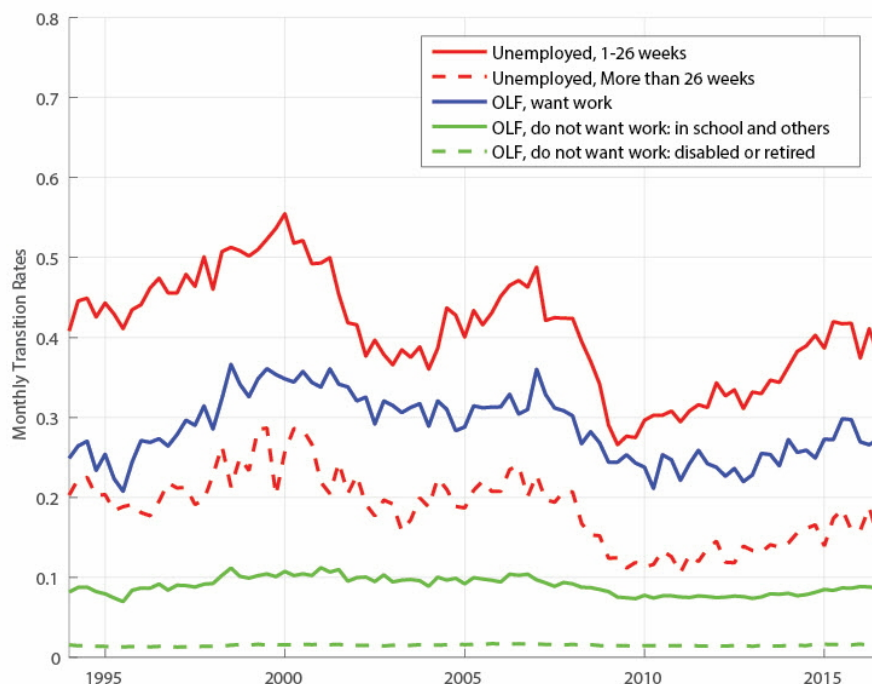
Heterogeneous Search Pools

We have motivated the NEI in Hornstein et al. (2014) as a way to capture persistent differences in labor market attachment across groups through their average employment transition rates. The same persistent differences in transitions to employment play an integral part in the generalized matching function with heterogeneous search efficiencies described in Section 1. From this perspective, the important

Figure 1 Labor Force Status (LFS) of the Nonemployed

distinctions between the different nonemployment states that enter the NEI and the generalized matching function are (1) short-term unemployment and (2) long-term unemployment, (3) those who are OLF and want to work, (4) those who are OLF, do not want to work, are in school, and others, and (5) those who are OLF, do not want to work, and are disabled or retired. For this aggregation of nonemployment states, the differences of employment transitions across groups clearly dominate the differences within groups. We now describe how the composition and the employment transitions of this “aggregated” search pool change with the business cycle.

In Figure 1, we plot the working-age population shares of the five aggregated nonemployment segments for the period 1994–2016. From this graph it is apparent that for the two recessions in the sample period, 2001 and 2007–09, the nonemployment share is increasing mainly because of increased unemployment. The increase of LTU in the Great Recession is especially striking. Following the recovery from the Great

Figure 2 LFS Contingent Transition Rates to Employment

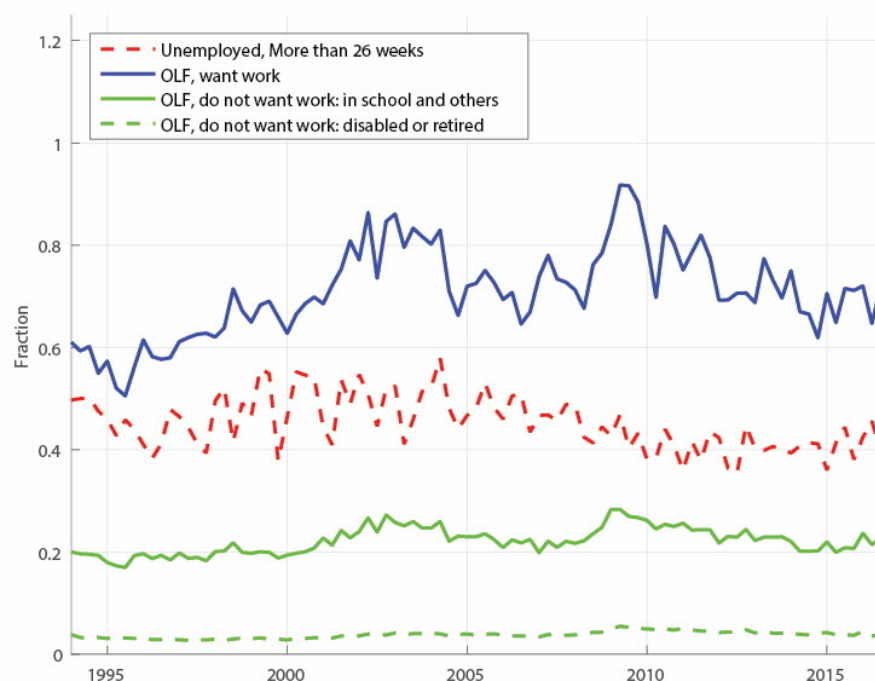
Recession, the decline in unemployment was compensated by an increase of those who are disabled or retired such that the working-age share of nonemployment remained elevated.

In Figure 2, we plot the employment transition rates of the five aggregated nonemployment segments for the period 1994–2016.⁹ The figure reflects the persistent differences in employment transition rates across different nonemployment segments. In particular, employment transition rates across nonemployment segments move together, they decline in recessions and increase in recoveries such that the ranking of transition rates remain unchanged.¹⁰ This does not preclude different cyclical sensitivities for the transition rates of different nonemployment

⁹ The “aggregated” employment transition rates are calculated as the nonemployed weighted averages of the employment transition rates calculated using data on exit probabilities to employment and other nonemployment states.

¹⁰ There also appears to be a secular decline in employment transition rates for unemployed and those OLF who want to work.

Figure 3 LFS Contingent Transition Rates Relative to Transition Rates of Short-Term Unemployed



segments, but it appears that the volatility of employment transition rates relative to those of the STU is limited, Figure 3.¹¹

In Table 3, we summarize the average properties of working-age population shares and relative employment transitions for our five aggregated nonemployment segments. As we have noted, nonemployment has somewhat increased in the years following the Great Recession, and most of the increase took place among the LTU and the disabled and retired, Table 3 column 1. Even though transitions to employment declined substantially following the Great Recession, the decline affected all nonemployment segments equally, such that the transitions of all segments relative to those of the STU remained quite stable. This stability of relative employment transitions holds independently of how we measure employment transitions, whether it is the straight

¹¹ Hornstein and Kudlyak (2016) use these different cyclical sensitivities to identify differences in search effort across segments.

Table 3 Aggregated Nonemployment Categories

	WAP Share		Relative Transition			
	(1a)	(1b)	Probability (2a)	Rate (2b)	Rate (3a)	Rate (3b)
Unemployed						
Short term	2.8	3.0	1.00	1.00	1.00	1.00
Long term	0.6	1.6	0.53	0.47	0.48	0.41
OLF, want to work						
Marg att and others	2.3	2.4	0.50	0.51	0.69	0.74
OLF, do not want to work						
In school and others	11.3	11.9	0.27	0.29	0.21	0.23
Disabled or retired	19.6	21.5	0.05	0.06	0.03	0.04

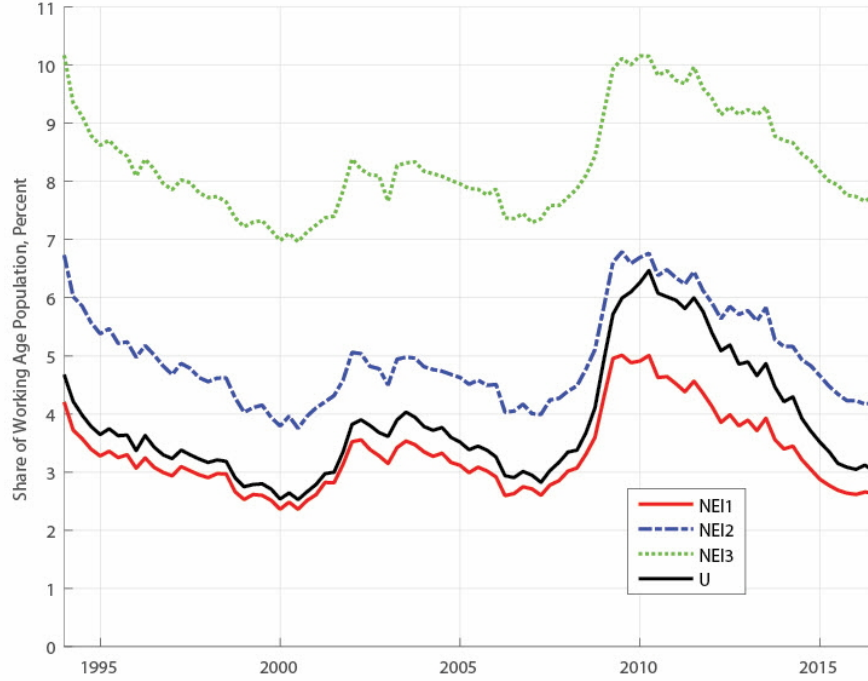
Note: For the different nonemployed population groups, columns 1 display their average percentage shares in total working-age population (WAP). Columns 2 display the average of their employment transition probabilities relative to the transition probabilities of the STU, and columns 3 display the average of their employment transition rates relative to the transition rates of the STU when the employment transition rates have been calculated using the exit probabilities to employment and different nonemployment states as described in the Appendix. Columns a cover the time period 1994–2007 and b the time period 2008–16.

employment transition probability, Table 3 column 2, or the employment transition rate calculated from the exit probabilities to employment and a different nonemployment state, Table 3 column 3. In Section 2, we have argued that the employment transition rate represents a better measure of employment transitions, and for the following, we will use the average employment transition rates for the full sample, the average of Table 3 column 3a and column 3b, as our measure of the relative quality of the different nonemployment segments.¹²

3. MATCHING EFFICIENCY AND THE NEI

We now use the information on relative employment transition rates to construct measures of quality-adjusted search input for a matching function with heterogeneous search efficiencies as described in Section 1, equation (4). These quality-adjusted search input measures correspond to the nonemployment index we proposed in Hornstein et al. (2014). We then show that measures of matching efficiency for generalized matching functions that account for heterogeneity are less volatile than the matching efficiency measures derived from standard

¹² Using average relative transition rates from the pre-2008 period does not change the results.

Figure 4 NEI: A Measure of Quality-Adjusted Search Effort

matching functions that assume homogeneous search and are limited to the unemployment pool.

We proceed by gradually expanding our definition of the search pool. For the first definition (NEI1), we take the weighted sum of STU and LTU, where STU receives a weight of 1 and LTU receives a weight of 0.46. The weight of LTU is its average employment transition rate relative to STU or, using the heterogeneous search framework

$$\frac{\lambda_{LTU}}{\lambda_{STU}} = \frac{\rho_{LTU}}{\rho_{STU}} = \rho_{LTU},$$

since $\rho_{STU} \equiv 1$.¹³ For the second definition (NEI2), we add the OLF who want to work with a weight of 0.71 to NEI1. Finally, for the third

¹³ Assigning a weight of one to STU is a normalization. Choosing a different weight for STU while maintaining the relative weights between the different groups affects the scale of the NEI but not its cyclical properties.

definition (NEI3), we add the OLF who are at school with a weight of 0.24 and the disabled and retired with a weight of 0.04 to NEI2. The working-age population shares of the three quality-adjusted search pools are displayed in Figure 4. For comparison, we have also added the working-age population share of the unweighted unemployed (U), which represents the standard measure of unemployment.

By construction, the level of the NEIs is increasing as we expand the coverage of nonemployment. In particular, once we include weighted OLF (NEI2 and NEI3), the levels of the NEIs are larger than for the standard measure of unemployment U. But note that the NEIs tend to be less volatile than the standard measure of unemployment, that is, they increase less in recessions than does the standard measure of unemployment. Furthermore, like the unemployment rate, all NEIs have essentially returned to their pre-Great Recession lows.

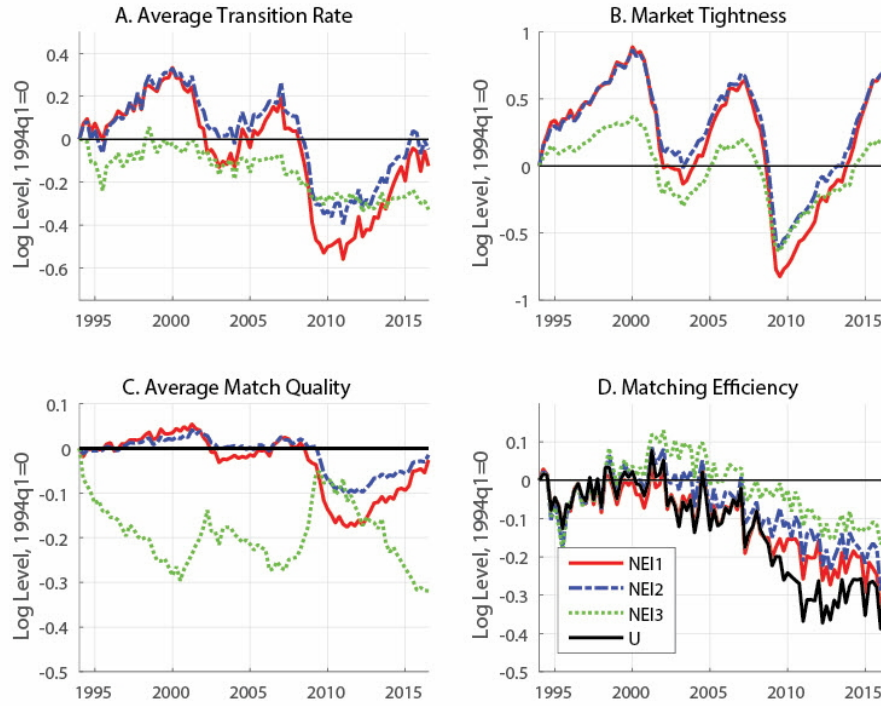
The proposed NEIs represent the quality-adjusted input to a generalized matching function that accounts for heterogeneity in search efficiencies across types. Following the discussion in Section 1, we can decompose changes in the average employment transition rate across all nonemployment segments included in an NEI, $\bar{\lambda}$, into changes coming from market tightness, θ , average search pool quality, $\bar{\rho}$, and aggregate matching efficiency, κ , equation (7). We construct market tightness, that is, the ratio of vacancies to the unweighted sum of nonemployment segments in the NEI, using the adjusted help-wanted index (HWI) from Barnichon (2010) for vacancies and posted job openings from JOLTS.¹⁴ In Figure 5, we plot the average employment transition rates (A), market tightness (B), average quality (C), and matching efficiency (D) for our three NEI definitions.¹⁵ For comparison, we also plot average quality and matching efficiency for the standard measure of unweighted unemployment.

The average employment transition rate declines in recessions and increases in expansions, Figure 5.A. This property of the average transition rate simply reflects the same countercyclical pattern for all of the component transition rates. As we expand the coverage of the search pool, the average transition rate becomes less volatile.¹⁶ In

¹⁴ The HWI index is available from the 1970s on, whereas JOLTS data are available only from 2000 on. The shift in job advertising from print media to web-based means that the HWI may not be consistent over time. Barnichon (2010) corrects for these structural changes in the HWI series in a way such that the HWI lines up with the JOLTS job openings in mid-2000, and we splice the two series in 2006.

¹⁵ We plot the log of each series and normalize each series to zero at the beginning of the sample.

¹⁶ The level of the average employment transition rate also declines as we expand the coverage of the search pool, but this is not apparent from Figure 5.A since we have normalized each series to zero at the beginning of the sample.

Figure 5 Components of the Average Employment Transition Rate

particular, the average transition rate declines less in recessions. This is because relative to the employment transition rates of the unemployed, the transition rates of the OLF (want work) decline less in recessions (NEI2 versus NEI1), as do the transition rates of the OLF (do not want work) (NEI3 versus NEI2). Furthermore, the unemployed with highly volatile transition rates represent a relatively small share of NEI3.

Market tightness has the same cyclical pattern as the average employment transition rate: it declines in recessions and increases in expansions, Figure 5.B. This feature reflects the fact that in recessions vacancy postings decline and nonemployment increases. The volatility of market tightness also declines as we expand the coverage of the search pool, and this reflects the fact that unweighted, like weighted, (NEI) nonemployment becomes less volatile as we expand the coverage of the search pool, Figure 4.

In the standard matching framework with homogeneous search, average quality is constant. In the generalized matching framework with heterogeneous search, average quality reflects the composition of the search pool, Figure 5.C. For example, average quality for quality-adjusted unemployment (NEI1) declines in recessions because the share of LTU with relatively low search efficiency is increasing in recessions. Average quality continues to decline in recessions for the search pool (NEI2) that includes OLF (want work), but the magnitude of the decline is reduced since the weight of OLF (want work) is more similar to STU than it is to LTU. For the broadest definition of the search pool (NEI3) that includes OLF (do not want work), average quality increases in recessions. This is unlike what we see for the two narrower definitions of the search pool and occurs because the share of OLF (do not want work) in total nonemployment declines in recessions and both components of OLF (do not want work) receive smaller quality weights than all other nonemployment components in the search pool.

Finally, matching efficiency represents the residual component that, together with market tightness and average quality, accounts for the movements in average employment transition rates. In Figure 5.D, we use equation (7) to construct measures of matching efficiency for the different search pool definitions. We assume that the matching elasticity is $\alpha = 0.35$, a value consistent with estimates from Barnichon and Figura (2015) and within the range of reported matching elasticities from Petrongolo and Pissarides (2001). We start with the matching efficiency calculated for the standard search pool definition with homogeneous unemployment (U). For this measure, the decline in matching-elasticity weighted market tightness accounts for some of the decline in average transition rates, but with no change in average quality a significant decline in matching efficiency remains. Once we account for heterogeneity in the search pool of unemployed (NEI1), average quality declines in recessions and less of a decline in matching efficiency is required. Once we include OLF (want work) in the search pool (NEI2), the average transition rate and market tightness both decline less in recessions, but the change is more pronounced for the average transition rate such that a smaller decline of matching efficiency is required. Finally, for the most comprehensive definition of the search pool (NEI3), which includes OLF (do not want work), average employment transition rates are even less volatile relative to market tightness and average quality increases in recessions such that substantially smaller declines in matching efficiency occur during recessions.

4. CONCLUSION

We have reviewed the evidence on heterogeneity among the nonemployed in the CPS with respect to their likelihood of making the transition to employment within a month, and we have shown that while the differences between the groups that are most and least likely to make the transition to employment are quantitatively substantial, there is also a gradual transition between the groups at the extremes. We have then shown that the NEI proposed in Hornstein et al. (2014) represents the quality-adjusted search input of a generalized matching function that accounts for heterogeneity in search efficiency across the search pool. Finally, expanding the coverage of the search pool at the same time one accounts for heterogeneity in search effort reduces the measured decline in matching efficiency associated with the Great Recession. In other words, for an appropriately defined broader concept of nonemployment, the efficiency of the U.S. labor market has not declined as much as would be suggested by standard measures of unemployment.

APPENDIX

Data for the population shares and employment transition rates for nonemployment by reason are constructed from the monthly CPS micro datasets as in Kudlyak and Lange (2014). All data are seasonally adjusted using the procedure proposed by Watson (1996). We deviate from Hornstein et al. (2014) in the construction of the employment transition rates in order to account for the possibility that the nonemployment state may change not only because a nonemployed worker makes the transition to employment, but also because she may just make the transition to a different nonemployment state. Both transition rates will be reflected in the transition probability to employment, but from a matching function perspective we are mainly interested in the transition rate to employment.

Take a group with nonemployment status j . Assume that transitions to employment or a different nonemployment state arrive continuously according to Poisson processes with arrival rates λ_{jE} and λ_{jN} , respectively. Then the probability that within a month a member will

exit nonemployment state j for employment is

$$p_{jE} = \int_0^1 \underbrace{e^{-\lambda_{jN}\tau}}_{\text{no exit to } N \text{ by } \tau} \underbrace{\left(\lambda_{jE}e^{-\lambda_{jE}\tau}\right)}_{\text{transition to } E \text{ at } \tau} d\tau = \lambda_{jE} \int_0^1 e^{-(\lambda_{jN}+\lambda_{jE})\tau} d\tau,$$

ignoring the possibility that somebody will flow back into state j in the same month.¹⁷ We can simplify this expression and apply the same procedure to the exit probability to a different nonemployment state, and we get

$$\begin{aligned} p_{jE} &= \frac{\lambda_{jE}}{\lambda_{jE} + \lambda_{jN}} \left[1 - e^{-(\lambda_{jE} + \lambda_{jN})} \right] \\ p_{jN} &= \frac{\lambda_{jN}}{\lambda_{jE} + \lambda_{jN}} \left[1 - e^{-(\lambda_{jE} + \lambda_{jN})} \right]. \end{aligned}$$

We have data on the monthly transition probabilities to employment, p_{jE} , or a different nonemployment state, p_{jN} . We can recover the transition rates λ from the transition probabilities p as follows

$$\begin{aligned} \lambda_{jN} &= -\frac{\log(1 - p_{jE} - p_{jN})}{(1 + p_{jE}/p_{jN})} \\ \lambda_{jE} &= -\frac{\log(1 - p_{jE} - p_{jN})}{(1 + p_{jE}/p_{jN})} \\ &= -p_{jE} \frac{\log(1 - p_{jE} - p_{jN})}{(p_{jE} + p_{jN})}. \end{aligned}$$

For p_{jN} small relative to p_{jE} we have

$$\lambda_{jE} \approx -\log(1 - p_{jE}),$$

that is, we can limit attention to the employment transition probabilities. Note that the exit rates are defined on the unit interval, which represents one month. So we are calculating monthly exit rates.

¹⁷ Shimer (2012) proposes a procedure that recovers continuous time exit rates allowing for the possibility that an agent who exits a state returns to the state within the unit of observation. His procedure uses information from the complete transition matrix covering transitions between all labor market states.

REFERENCES

- Appelbaum, Binyamin. 2014. "Still Needed: Millions of Jobs." *New York Times*, April 4.
- Barnichon, Regis. 2010. "Building a Composite Help-Wanted Index." *Economics Letters* 109 (December): 175–78.
- Barnichon, Regis, and Andrew Figura. 2015. "Labor Market Heterogeneity and the Aggregate Matching Function." *American Economic Journal: Macroeconomics* 7 (October): 222–49.
- Diamond, Peter. 2013. "Cyclical Unemployment, Structural Unemployment." *IMF Economic Review* 61 (August): 410–55.
- Elsby, Michael W.L., Bart Hobijn, and Ayşegül Şahin. 2015. "On the Importance of the Participation Margin for Labor Market Fluctuations." *Journal of Monetary Economics* 72 (May): 64–82.
- Flinn, Christopher J., and James J. Heckman. 1983. "Are Unemployment and Out of the Labor Force Behaviorally Distinct Labor Force States?" *Journal of Labor Economics* 1 (January): 28–42.
- Fujita, Shigeru. 2014. "On the Causes of Declines in the Labor Force Participation Rate." Federal Reserve Bank of Philadelphia *Research Rap Special Report* (February).
- Hornstein, Andreas, and Marianna Kudlyak. 2016. "Estimating Matching Efficiency with Variable Search Effort." Federal Reserve Bank of Richmond Working Paper 16-13R (December).
- Hornstein, Andreas, Marianna Kudlyak, and Fabian Lange. 2014. "Measuring Resource Utilization in the Labor Market." Federal Reserve Bank of Richmond *Economic Quarterly* 100 (First Quarter), 1–21.
- Hornstein, Andreas, Marianna Kudlyak, Fabian Lange, and Tim Sablik. 2014. "Does the Unemployment Rate Really Overstate Labor Market Recovery?" Federal Reserve Bank of Richmond *Economic Brief* 14-06 (June).
- Irwin, Neil. 2017. "Trump Said the Unemployment Rate Wasn't Real. Here Are Some Other Options." *New York Times*. February 3.
- Kudlyak, Marianna, and Fabian Lange. 2014. "Measuring Heterogeneity in Job Finding Rates among the Nonemployed Using Labor Force Status Histories." Federal Reserve Bank of Richmond Working Paper 14-18 (October).

- Madrian, Brigitte C., and Lars John Lefgren. 1999. "A Note on Longitudinally Matching Current Population Survey (CPS) Respondents." Cambridge, Mass.: National Bureau of Economic Research Technical Working Paper 247 (November).
- Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (June): 390–431.
- Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory, 2nd edition*. Cambridge, Mass.: MIT Press.
- Polivka, Anne E., and Stephen M. Miller. 1998. "The CPS After the Redesign: Refocusing the Economic Lens." In *Labor Statistics Measurement Issues; Studies in Income and Wealth*, vol. 60, edited by John Haltiwanger, Marilyn E. Manser, and Robert Topel. Chicago: University of Chicago Press, 249–89.
- Shimer, Robert. 2012. "Reassessing the Ins and Outs of Unemployment." *Review of Economic Dynamics* 15 (April): 127–48.
- Veracierto, Marcelo. 2011. "Worker Flows and Matching Efficiency." Federal Reserve Bank of Chicago *Economic Perspectives* 35 (Fourth Quarter): 147–69.
- Watson, Mark W. 1996. "Comment on 'Is Seasonal Adjustment a Linear or Nonlinear Data-Filtering Process.'" *Journal of Business and Economic Statistics* 14 (July): 394–6.
- Yellen, Janet L. 2014. "Labor Market Dynamics and Monetary Policy." Speech at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyo., August 22.

Price Dispersion When Stores Sell Multiple Goods

Nicholas Trachter

Search frictions are a prominent departure from the standard style of model we tend to write, which relies on frictionless Walrasian markets. They are not only prominent because they help us construct interesting models where policy can play a particularly important role, but also because search frictions are relatively easy to measure in the data. A large fraction of the literature on search frictions dwells with models of product markets where, for one reason or another, customers face a cost to act in the market (i.e., pay a search or switching cost to switch stores, pay a cost to learn a set of prices, etc.). A well-known result in a large class of models (based on the seminal work of Burdett and Judd [1983]) is that price dispersion for identical goods arises in equilibrium.

The empirical evidence on price dispersion for product markets — a good measure of the extent of the friction, as there should be no price dispersion for homogeneous goods in a Walrasian market — is large, mostly documenting dispersion for particular goods in retail markets. The literature abstracts from several important features of retail markets. One of these features is that most stores sell multiple goods, a feature that not only changes the measurement of search frictions, but also opens new avenues for theoretical research, given the scant availability of models of multiproduct pricing, i.e., models where firms price multiple goods simultaneously. In this paper, I review the work of Kaplan et al. (2016) (KMRT from now on), which is a recent study on the empirical properties of price dispersion in a multiproduct setting and provides a model to rationalize it.

■ The views expressed in this article are those of the author and do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: nicholas.trachter@rich.frb.org.

Most models of price dispersion feature retailers selling a single good. Thus, claims about price dispersion across goods are also claims about dispersion in prices across retailers. However, this correlation across stores and goods does not need to be perfect, for example, if the choice of a price of an individual good is not independent of a retailer's choices of prices for any other goods sold at his store. In fact, if stores sell multiple goods, we can understand whether dispersion arises at the store level or if dispersion arises at the store-good level. Exploring the forces driving price dispersion lets us understand the frictions we need to introduce into our models.

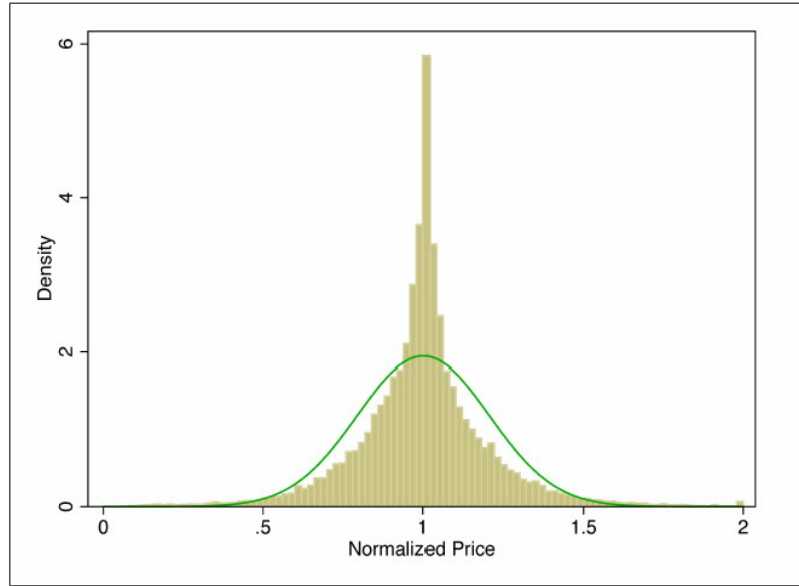
KMRT attempts to provide answers to the origins of price dispersion. Empirically, it does so by exploiting some recently available large-scale datasets. The Kilts-Nielsen Retail Scanner (KNRS) dataset provides an ideal laboratory to study price dispersion with multiproduct retailers (i.e., retailers that sell multiple goods). The KNRS provides weekly price and quantity information for around 1.5 million goods — a good is defined by its Universal Product Code or UPC — at about 40,000 stores across the United States from 2006 to 2012. The vast amount of information in datasets like the KNRS allows researchers to provide novel insights to the measurement of price dispersion. KMRT finds that there is a large amount of price dispersion for identical goods — standard deviation of 15 percent — and that a large part of this dispersion is due to stores with the same average price level pricing individual goods in persistently different ways. This finding, not shown before in the literature, is coined by the authors as *relative price dispersion*. A similar feature was found by Gorodnichenko et al. (2015) for stores selling multiple goods in online markets.

In this paper, I review the basics of the empirical findings of KMRT regarding relative price dispersion, and I also provide a review of the basics of the theoretical model the authors develop to explain their empirical findings. The paper is full of robustness exercises (for the empirical analysis) and validation exercises (for the main mechanism that the paper puts forward). The objective of this paper is to introduce the reader to this exciting avenue for research.

1. RELATIVE PRICE DISPERSION IN THE DATA

Let p_{jst} denote the price of good $j = 1, 2, \dots, J$ at store $s = 1, 2, \dots, S$ at week t . To make goods comparable (i.e., butter is much cheaper than caviar) it is useful to normalize all prices. With this in mind, let

$$\hat{p}_{jst} = \ln p_{jst} - \frac{\sum_s \ln p_{jst}}{S}$$

Figure 1 Distribution of Normalized Prices

denote the normalized price of a given good in a particular geographical region.¹ The value \hat{p}_{jst} measures the (log) relative price of a good j sold by store s relative to the price of that good sold by every store in the geographical region, at week t . For example, if $\hat{p}_{jst} = 0.1$, we have that, at time t , good j is 10 percent more expensive in store s than in the other stores in the area. Likewise, when $\hat{p}_{jst} = -0.1$, we have that the good is 10 percent cheaper at store s .

Figure 1 plots the average distribution of normalized prices across all goods, markets, and time periods (the distribution is expenditure weighted), borrowed from Kaplan and Menzio (2015), which uses data from the KNRS dataset. Also, to aid in the analysis, the figure plots the density of a normal distribution with the same mean and variance. As it can be seen, the price distribution exhibits higher kurtosis, with a high concentration of mass close to the mean. More importantly, price dispersion is large, with a standard deviation for normalized prices, \hat{p}_{jst} , of 0.15.

¹ The boundaries of the region define the set of stores to be included and thus define the set S .

What explains the extent of price dispersion we observe in the data? How much of the price dispersion that we observe comes from the fact that different stores have different price levels (store component)? How much comes from the fact that stores price the different goods they sell in different ways (store-good component)? How much is transitory, and how much is persistent? Campbell and Eden (2014) noted that for a subsample of the KNRS, the store component does not explain all of the variation. In other words, they noted that some of the variability needs to come from the store-good component. Lewis (2008) observed something similar for the price of the same kind of gasoline at different gas stations. With the aim to decompose price dispersion, we can write the price of good j at store s at week t as

$$\hat{p}_{jst} = \hat{y}_{st} + \hat{z}_{jst} .$$

The term \hat{y}_{st} accounts for the store component (i.e., the price level of the store) and is defined as $\hat{y}_{st} = \sum_j \hat{p}_{jst}/J$. The term \hat{z}_{jst} is the store-good component, and it is defined as a residual: $\hat{z}_{jst} = \hat{p}_{jst} - \hat{y}_{st}$. The store component captures the extent to which a store tends to be more expensive than other stores, regardless of each individual good that it sells, while the store-good component captures variation in relative prices across goods for a particular store.

Furthermore, a statistical model can be posed for each component (i.e. the store and store-good components) in order to understand their persistence. A particularly appealing model is to use an ARMA(1,1) representation for each component, with the intention of capturing persistent variation with the autoregressive component and transitory variations with the moving average component. Table 1 presents the variance decomposition for the baseline scenario considered in KMRT, which is restricted to the Minneapolis-St. Paul Designated Market Area (DMA), which is roughly consistent with the Minneapolis-St. Paul Metropolitan Statistical Area (MSA). Also, the baseline scenario restricts the analysis to include only 1,000 goods — those with the highest revenue in the DMA.

As the table shows, the standard deviation of normalized prices is 0.153. The standard deviation of the store component is 0.06, and the standard deviation of the store-good component is 0.141. In fact, the variance decomposition implies that only 15.5 percent of the variation of prices is explained by the store component, while the rest — 84.5 percent of variance — is explained by the store-good component. On the one hand, the relatively low importance of the store component implies that explanations for price dispersion that follow from store differentials are not that relevant. Standard explanations of the store component

Table 1 Dispersion in Prices: Persistent and Transitory (from KMRT)

	Variance	Percent	Standard Deviation
Store component			
Transitory	0.000	3.2	0.011
Fixed plus persistent	0.004	96.8	0.059
Total Store	0.004	100.0	0.060
Store-good component			
Transitory	0.013	64.1	0.113
Fixed plus persistent	0.007	35.9	0.084
Total store-good	0.020	100.0	0.141
Total	0.023	100.0	0.153

Note: The left column presents the cross-sectional variances of UPC prices, as well as the store and store-good components separately. The middle columns present the decomposition of this variance into persistent and transitory components. The right column presents the cross-sectional standard deviations.

are those that stem from heterogeneous cost structures across stores and heterogeneity across stores with respect to the amenities provided to shoppers (i.e., differentials in the shopping experience that can be translated into price differentials). On the other hand, the relatively high importance of the store-good component implies that we need to focus our attention on theories that explain why stores with the same overall price level price individual goods in different ways.

Around 65 percent of the variance of the store-good component is explained by its transitory components, while 35 percent of the variance is explained by highly persistent components. The literature offers compelling theories of transitory differences in the price of the same good across equally expensive stores. For instance, according to the theory of intertemporal price discrimination (see, e.g., Conlisk et al. 1984; Sobel 1984; and Menzio and Trachter 2015a), sellers find it optimal to occasionally lower the price of a particular good in order to discriminate between low-valuation customers who are willing to do their shopping at any time during the month and high-valuation customers who need to make their purchases on a specific day of the month. As different sellers implement these occasional price reductions at different times, the equilibrium may feature short-term differences in the price of the same good across equally expensive stores. According to the inventory management theory (see, e.g., Aguirregabiria 1999), a seller finds it optimal to increase the price of a good as the inventory of the good falls and to lower the price when the inventory of the good is replenished. As different sellers have different inventory cycles, the

Table 2 Robustness (from KMRT)

	Low price		High price		Low durability		High durability	
	Sd	Dec/ %	Sd	Dec/ %	Sd	Dec/ %	Sd	Dec/ %
Store								
Transitory	0.024	8.7	0.025	15.6	0.013	4.0	0.027	27.9
Fixed plus persistent	0.078	91.3	0.059	84.4	0.062	96.0	0.043	72.1
Total Store	0.082	20.6	0.065	15.9	0.063	19.3	0.051	19.4
Store-good								
Transitory	0.122	57.4	0.130	77.0	0.103	64.0	0.077	55.8
Fixed plus persistent	0.105	42.6	0.071	23.0	0.077	36.0	0.069	44.2
Total store-good	0.161	79.4	0.148	84.1	0.129	80.7	0.103	80.6
	Unilever		Coca-Cola		State: MN		County: Hennepin	
	Sd	Dec/ %	Sd	Dec/ %	Sd	Dec/ %	Sd	Dec/ %
Store								
Transitory	0.035	27.4	0.030	15.5	0.011	2.5	0.015	6.2
Fixed plus persistent	0.058	72.6	0.070	84.5	0.070	97.5	0.058	93.8
Total Store	0.068	21.3	0.076	26.2	0.071	17.6	0.060	12.5
Store-good								
Transitory	0.101	60.9	0.106	68.9	0.120	60.9	0.128	64.4
Fixed plus persistent	0.081	39.1	0.071	31.1	0.096	39.1	0.095	35.6
Total store-good	0.130	78.7	0.127	73.8	0.154	82.4	0.159	87.5

Note: This table presents a set of robustness exercises developed in KMRT. In particular: the low- and high-price samples, the low- and high-durability samples, the Unilever and Coca-Cola samples, and alternative definitions of a market (state of Minnesota and Hennepin County).

equilibrium may feature short-term differences in the price of the same good across equally expensive stores. However, little has been made in the literature to understand the persistent component that, following KMRT, I will describe as relative price dispersion. Before moving to the description of a simple theory of relative price dispersion, I want to discuss some of the robustness exercises in terms of the variance decomposition results. These exercises will shed light on why existing theories cannot explain relative price dispersion. The robustness exercises are provided in Table 2.

High- and low-price goods. A potential explanation for relative price dispersion is managerial inattention (Ellison et al. 2015). According to this story, equally expensive stores may set persistently different

prices for the same good because managers choose to not pay much attention to the price of low-ticket items. With this in mind, KMRT looks at relative price dispersion for low-price and high-price goods. The low-price subsample features more relative price dispersion than the full sample: the store-good component accounts for 79 percent of the overall variance of prices, of which the persistent components account for 43 percent. The high-price subsample features less relative price dispersion than the full sample, but relative price dispersion is still a substantial fraction of overall price dispersion. Hence, relative price dispersion is not only a feature of low-price, low-revenue goods and thus is unlikely to be entirely due to managerial inattention.

Goods from a single distributor. Another possible explanation for relative price dispersion is that equally expensive stores set persistently different prices for the same good because they have better or worse relationships (and, hence, are charged lower or higher prices) with the wholesaler. With this in mind, KMRT decomposes price dispersion for a subset of products produced and distributed by a single wholesaler. If relative price dispersion is caused by different retailer-wholesaler relationships, relative price dispersion should be absorbed by the store component when we restrict attention to products from a single wholesaler. The paper considers two subsamples of goods: goods produced by Coca-Cola and by Unilever. For both samples of goods, the overall degree of price dispersion is very similar to the degree of price dispersion in the baseline sample. However, the fraction of variation that is due to the store component is somewhat larger: 21 percent for Unilever and 26 percent for Coca-Cola, compared with 16 percent for the baseline. This is consistent with the hypothesis that some part of price dispersion is due to different relationships between particular stores and particular distributors. However, for both of these distributors, the vast majority of price dispersion is due to the store-good component, and, of this, the persistent parts account for 39 percent (Unilever) and 31 percent (Coca-Cola). Thus, relative price dispersion exists even when only considering goods from the same distributor and so is not only driven by heterogeneity in distributional relationships.

Low- and high-durability goods. Another natural explanation for relative price dispersion is shelf management. Some stores may keep perishable goods on their shelves for longer and, for this reason, sell them at systematically lower prices, while other stores may remove perishable goods sooner and, for this reason, sell them at systematically higher prices. To evaluate this story, KMRT decomposes price dispersion separately for two subsamples of goods: low-durability goods (i.e., perishable goods) and high-durability goods. Even though the two subsamples contain very different sets of products, the overall

decomposition of price dispersion is quite similar. For both subsamples, the store component accounts for approximately 20 percent and the store-good component for 80 percent of the cross-sectional variance of prices. For both subsamples, the transitory part accounts for roughly two-thirds and the persistent part for roughly one-third of the cross-sectional variance of the store-good component of prices. These findings suggest that relative price dispersion is unlikely to be a phenomenon caused by different styles of shelf management for perishable goods. Indeed, relative price dispersion turns out to be slightly more important in the subsample of goods that are less perishable.

Markets. The baseline analysis focused on a single geographic region, the Minneapolis-St. Paul DMA. To show that the results do not depend on the particular level of geographic aggregation, Table 2 also considers alternative levels of geographic aggregation for the definition of a market. In particular, it reports the variance decomposition when we use a broader definition of market (the state of Minnesota) and a narrower definition of a market (Hennepin County, which is contained in the Minneapolis-St. Paul DMA). All findings are robust to switching to either of these alternative levels of aggregation.

2. A MODEL OF RELATIVE PRICE DISPERSION

In this section, I consider the model developed and used in KMRT to explain the concept of relative price dispersion. The model is a variation of Burdett and Judd (1983), which is the workhorse model to explain equilibrium price dispersion across stores selling a single homogeneous good. In KMRT, the model is extended to allow for multiple goods (in particular, two goods) and to allow for heterogeneity in customer shopping behavior. The latter assumption follows from the observation in the data that there is heterogeneity in the number of stores that customers visit. This assumption is critical in order to obtain relative price dispersion.

Consider a market populated by homogeneous sellers and heterogeneous buyers who trade two goods (i.e., good 1 and good 2). Specifically, the market is populated by a measure $s > 0$ of identical sellers. Every seller is able to produce each of the two goods at the same constant marginal cost, normalized to zero. Every seller chooses a price for good 1, p_1 , and a price for good 2, p_2 , so as to maximize his profits, taking as given the distribution $H(p_1, p_2)$ of the vector of prices across sellers. Denote as $F_i(p)$ the fraction of sellers whose price for good $i \in \{1, 2\}$ is smaller than p . Here, $F_i(p)$ refers to the distribution of prices for good $i \in \{1, 2\}$. Similarly, let $G(q)$ denote the fraction of

sellers whose prices p_1 and p_2 sum up to less than q . $G(q)$ refers to the distribution of basket prices.

On the other side of the retail market, there is a measure 1 of buyers. A fraction $\mu_b \in (0, 1)$ of buyers are of type b and a fraction $\mu_c = 1 - \mu_b$ of buyers are of type c , where b stands for busy and c stands for cool. A buyer of type b demands one unit of each good, for which he has valuation $u_b > 0$. A buyer of type c demands one unit of each good, for which he has valuation u_c , with $u_b > u_c > 0$. More specifically, if a buyer of type $i \in \{b, c\}$ purchases both goods at the prices p_1 and p_2 , he attains a utility of $2u_i - p_1 - p_2$. If a buyer of type $i \in \{b, c\}$ purchases one of the two goods at the price p , he attains a utility of $u_i - p$. If a buyer of type $i \in \{b, c\}$ does not purchase any of the goods, he attains a utility of zero.

In the retail market, trade is frictional. Buyers cannot purchase from just any seller in the market, as each buyer only has access to a small network of sellers. In particular, a buyer of type b can access only one seller with probability $\alpha \in (0, 1)$ and two sellers with probability $1 - \alpha$. Similarly, a buyer of type c can access only one seller with probability α and two sellers with probability $1 - \alpha$. A buyer who can only access one seller is referred to as a captive buyer, and a buyer who can access multiple sellers is referred to as a noncaptive buyer. The authors interpret these restrictions on the buyers' access to sellers as physical constraints (i.e., sellers the buyer can easily reach) rather than as informational constraints (i.e., sellers of which the buyer is aware). Moreover, it is assumed that a buyer of type b must always make all of his purchases from just one of the sellers in his network. In contrast, a buyer of type c can purchase different goods from different sellers in his network. Again, the authors interpret this assumption as heterogeneity in the buyer's ability or willingness to visit multiple stores when shopping.

Notice that the model is static, as in Burdett and Judd (1983). The equilibrium price distribution resulting from the model should be interpreted as a long-term outcome. Indeed, in a repeated version of the model, it can be seen immediately that sellers would have nothing to gain from changing their prices over time. Moreover, in the presence of any type of adjustment costs, sellers would face a loss from changing their prices over time. Thus, in a repeated version of the model, sellers would keep their prices constant. Then, under this interpretation of the model, we should compare the equilibrium price distribution to the distribution of the persistent component of sellers' prices.

Equilibrium with Relative Price Dispersion

Consider an equilibrium in which some sellers have a basket price q greater than $u_b + u_c$ and some sellers have a basket price smaller than $u_b + u_c$ and greater than $2u_c$. Those sellers pricing a basket above $u_b + u_c$ will only sell baskets to busy shoppers, while those sellers pricing between $2u_c$ and $u_b + u_c$ will sell baskets to busy shoppers and one good to cool shoppers. KMRT refers to this type of equilibrium as a discrimination equilibrium, as in this equilibrium some sellers set their prices so as to discriminate between the high-valuation buyers who must purchase all the goods in the same location and the low-valuation buyers who can purchase different goods in different locations.

Sellers pricing baskets above $u_b + u_c$. Notice that it is not optimal for any seller in this region to set the price of either individual good above u_b .² Then, because no price is strictly above u_b also no price is equal or below u_c . As a result, sellers in this region do not sell goods to cool shoppers. Moreover, because no price is above u_b , the price of the basket $q = p_1 + p_2$ is below $2u_b$. Then, busy shoppers buy the basket of goods at these sellers. Because in this region only busy shoppers buy, and because they buy the basket of goods at price q , any combination of prices for good 1 and good 2 that give the same basket price q gives the same profits to the seller. Then, in this region, there will be indeterminacy of prices of good 1 and good 2, and the equilibrium will pin down the distribution of basket prices.

In this region, the profits of a seller are given by

$$S_1(q) = \mu_b[\alpha + 2(1 - \alpha)(1 - G(q))]q .$$

The seller is in the network of $\mu_b\alpha$ captive buyers of type b . A captive buyer of type b purchases both goods from the seller with probability 1, since $q < 2u_b$. The seller is also in the network of $\mu_b2(1 - \alpha)$ noncaptive buyers of type b . A noncaptive buyer of type b purchases both goods from the seller with probability $1 - G(q)$, which is the probability that the second seller in the buyer's network has a basket price greater than q . Finally, the seller is in the network of some buyers of type c , but these buyers do not buy from this seller.

The highest basket price, q_h , on the support of G equals $2u_b$. To see why, suppose that q_h is strictly smaller than $2u_b$. In this case, the profit for a seller with a basket price of q_h is then equal to $\mu_b\alpha q_h$, as this seller is the one with the highest basket price in the economy and, hence, only sells to captive buyers of type b . However, if the seller sets a basket price of $2u_b$, he attains a profit of $\mu_b\alpha 2u_b$, as the seller still

² To show this, it suffices to show that if a seller prices a good above u_b , there is a deviation to price at u_b that increases profits.

only sells to captive buyers of type b . Since $\mu_b \alpha q_h < \mu_b \alpha 2u_b$, it follows that the seller with a basket price of q_h is not maximizing his profit, and, hence, this cannot be an equilibrium. Hence, $q_h = 2u_b$.

Second, the support of G in this region is an interval $[q^*, q_h]$. To see why, suppose that the support of G has a gap between the basket price q_0 and the basket price q_1 . In this case, a seller with a basket price of q_0 attains a profit of $\mu_b[\alpha + 2(1 - \alpha)(1 - G(q_0))]q_0$. A seller with a basket price of q_1 attains a profit of $\mu_b[\alpha + 2(1 - \alpha)(1 - G(q_1))]q_1$. Since G has a gap between q_0 and q_1 , $G(q_0) = G(q_1)$ and the seller with a basket price of q_0 makes the same number of trades as a seller with a basket price of q_1 but enjoys a lower profit per trade. Therefore, the seller with a basket price of q_0 does not maximize his profit, and, hence, this cannot be an equilibrium.

It is now possible to solve for the distribution G in this region. At any point in the support of G it has to be the case that sellers attain the same profit. That is, $S_1(q) = S^*$. We can obtain S^* by evaluating $S_1(q)$ at $q^* = 2u_b$, which provides that $S^* = \mu_b \alpha 2u_b$ (given that $G(2u_b) = 1$). Then, we have that

$$\mu_b[\alpha + 2(1 - \alpha)(1 - G(q))]q = \mu_b \alpha 2u_b \text{ for all } q \in [q^*, 2u_b] .$$

Solving this equation with respect to $G(q)$ provides an expression for the equilibrium distribution of basket prices above $u_b + u_c$,

$$G(q) = 1 - \frac{\alpha}{2(1 - \alpha)} \frac{2u_b - q}{q} \text{ for } q \in [q^*, 2u_b] . \quad (1)$$

Sellers pricing baskets between $2u_c$ and $u_b + u_c$. As it happened for sellers pricing above $u_b + u_c$, no seller would choose here to price individual goods above u_b . Because of this, and because the basket price q of any seller in this region satisfies $2u_c < q \leq u_b + u_c$, we have that in this region sellers price one good below u_c and one good between u_c and u_b . As a result, sellers in this region sell baskets to busy shoppers and one good to cool shoppers. Say that the cheap good that the seller sells to the busy shopper is good i . Then, the profit of a seller in this region is given by

$$\begin{aligned} S_{2i}(q, p_i) &= \mu_b[\alpha + 2(1 - \alpha)(1 - G(q))]q \\ &\quad + \mu_c[\alpha + 2(1 - \alpha)(1 - F_i(p_i))]p_i . \end{aligned}$$

Even though we will not show it here, this expression makes use of the fact that $G(q)$ does not have mass points and $F_i(p)$ does not have mass points over the interval $(0, u_c]$.

An important result is that, for all $p \in [0, u_c]$, the fraction of sellers charging less than p for good 1 is exactly the same as the fraction of sellers charging less than p for good 2. That is, $F_1(p) = F_2(p) = F(p)$

for all $p \in [0, u_c]$. Because of this, the profit of a seller pricing in this region is symmetric in the two goods. That is,

$$S_{21}(q, p) = S_{22}(q, p) = S_2(q, p) .$$

Although I will not provide a proof here, the idea is intuitive. If $F_1(p) > F_2(p)$ for $p \in (p_0, p_1)$, with $0 \leq p_0 < p_1 \leq u_c$, then a seller posting the prices $(p, q - p)$ in this region would be better off posting the prices $(q - p, p)$ instead. In fact, the seller trades the basket of goods to the same number of type b buyers and at the same price by posting either $(q - p, p)$ or $(p, q - p)$. However, by posting $(q - p, p)$ rather than $(p, q - p)$, the seller trades the cheaper good to more type c buyers even though he charges the same price for it. Hence, if $F_1(p) > F_2(p)$ for $p \in (p_0, p_1)$, all sellers posting the prices $(p, q - p)$ in this region would be better off switching the price tags of the two goods until $F_1(p) = F_2(p)$.

A key result is that the profit of a seller pricing in this region attains its maximum at S^* for all basket prices q and prices of the cheaper good p such that q is in the interval $[q_l, u_b + u_c]$ and p is in the interval $[p_l, u_c]$, where q_l denotes the lower bound on the support of the price distribution of baskets and p_l denotes the lower bound on the support of the price distribution of an individual good. That is, $S_2(q, p) = S^*$ for all (q, p) such that $q \in [q_l, u_b + u_c]$ and $p \in [p_l, u_c]$. The proof of the statement is available in KMRT and follows the same strategy used in Menzio and Trachter (2015a). The idea of the proof is to show that if profits are not constant for all (q, p) such that $q \in [q_l, u_b + u_c]$ and $p \in [p_l, u_c]$, there are either gaps in the support of the distribution of G over the interval $[q_l, u_b + u_c]$ or gaps in the support of the distribution F over the interval $[p_l, u_c]$. In turn, if there are gaps in the support of one of the two distributions, there are some sellers who could increase their profits by either increasing the price of the basket or by increasing the price of one of the cheaper good.

We can now solve for the lowest basket price q^* posted by sellers pricing baskets above $u_b + u_c$, for the marginal distribution $G(q)$ for sellers pricing baskets below $u_b + u_c$, and for the marginal distribution $F(p)$ of prices among sellers in this region. Using that profits are maximized at S^* , and given that it has to be the case that $S_2(q, p) = S^*$ for $q \in [q_l, u_b + u_c]$ and $p \in [p_l, u_c]$, we can use $S_2(u_b + u_c, u_c) = S^*$ to obtain

$$\mu_b[\alpha + 2(1 - \alpha)(1 - G(u_b + u_c))](u_b + u_c) + \mu_c[\alpha + 2(1 - \alpha)(1 - F(u_c))]u_c = S^* . \quad (2)$$

Similarly, for a seller pricing a basket at q^* (recall that $q^* > u_b + u_c$) with both individual prices strictly above u_c and below u_b , it is also the case that attains the maximized profit S^* ,

$$\mu_b[\alpha + 2(1 - \alpha)(1 - G(q^*))]q^* = S^* . \quad (3)$$

Notice that the fraction of sellers with a basket price smaller than q^* is the same as the fraction of sellers with a basket price smaller than $u_b + u_c$, i.e., $G(q^*) = G(u_b + u_c)$. Also, notice that the fraction of sellers who charge less than u_c for good 1 is half of the fraction of sellers with a basket price smaller than q^* , i.e., $F(u_c) = G(q^*)/2$. Using these two observations together with equation (2) and equation (3) provides

$$\begin{aligned} & \mu_b[\alpha + 2(1 - \alpha)(1 - G(u_b + u_c))](u_b + u_c) + \\ & \mu_c[\alpha + 2(1 - \alpha)(1 - G(q^*)/2)]u_c \\ & = \mu_b[\alpha + 2(1 - \alpha)(1 - G(q^*))]q^* . \end{aligned}$$

We can solve this equation to find an expression for q^* (using equation (1) to obtain $G(q^*)$),

$$q^* = \frac{2\alpha(1 + u_c/u_b) + \alpha(\mu_c/\mu_b)(u_c/u_b)}{4\alpha - (2 - \alpha)(\mu_c/\mu_b)(u_c/u_b)} 2u_b . \quad (4)$$

We can use the fact that we figured out that profits are constant for all (q, p) such that $q \in [q_l, u_b + u_c]$ and $p \in [p_l, u_c]$ to obtain an expression for $G(q)$. Notice that a seller posting prices (p_1, p_2) such that $p_2 \in (u_c, u_b]$ and $q = p_1 + p_2 \in [q_l, u_b + u_c]$ attains the same profit as a seller posting prices (u_c, u_b) ,

$$\begin{aligned} & \mu_b[\alpha + 2(1 - \alpha)(1 - G(q))]q + \mu_c[\alpha + 2(1 - \alpha)(1 - F(u_c))]u_c \\ & = \mu_b[\alpha + 2(1 - \alpha)(1 - G(u_b + u_c))](u_b + u_c) + \\ & \quad \mu_c[\alpha + 2(1 - \alpha)(1 - F(u_c))]u_c . \end{aligned}$$

Using that $G(u_b + u_c) = G(q^*)$, we can solve this last equation to obtain an expression for the distribution of basket prices for $q \in [q_l, u_b + u_c]$,

$$G(q) = G(q^*) - \frac{\alpha + 2(1 - \alpha)(1 - G(q^*))}{2(1 - \alpha)} \frac{u_b + u_c - q}{q} \text{ for } q \in [q_l, u_b + u_c] . \quad (5)$$

Solving the equation $G(q_l) = 0$ with respect to q_l , we find that the lowest price on the support of the distribution of basket prices is given by

$$q_l = \frac{2\alpha u_b}{2 - \alpha} \frac{u_b + u_c}{q^*} . \quad (6)$$

Following the same argument as before, a seller posting prices (p_1, p_2) such that $p_1 \in [p_l, u_c]$, $p_2 \in (u_c, u_b]$, and $p_1 + p_2 = q_l$ attains the same profit as a seller posting prices $(u_c, q_l - u_c)$, i.e.,

$$\begin{aligned} & \mu_b[\alpha + 2(1 - \alpha)]q_l + \mu_c[\alpha + 2(1 - \alpha)(1 - F(p))]p \\ & = \mu_b[\alpha + 2(1 - \alpha)]q_l + \mu_c[\alpha + 2(1 - \alpha)(1 - F(u_c))]u_c . \end{aligned}$$

Again, using the fact that $F(u_c) = G(q^*)/2$ and solving the equation with respect to $F(p)$, we find that the distribution of good 1 prices for

$p \in [p_l, u_c]$ is given by

$$F(p) = \frac{G(q^*)}{2} - \frac{\alpha + 2(1 - \alpha)(1 - G(q^*)/2)}{2(1 - \alpha)} \frac{u_c - p}{p}. \quad (7)$$

Solving the equation $F(p_l) = 0$ with respect to p_l provides an expression for the lowest price on the support of the distribution of good 1 prices, which is given by

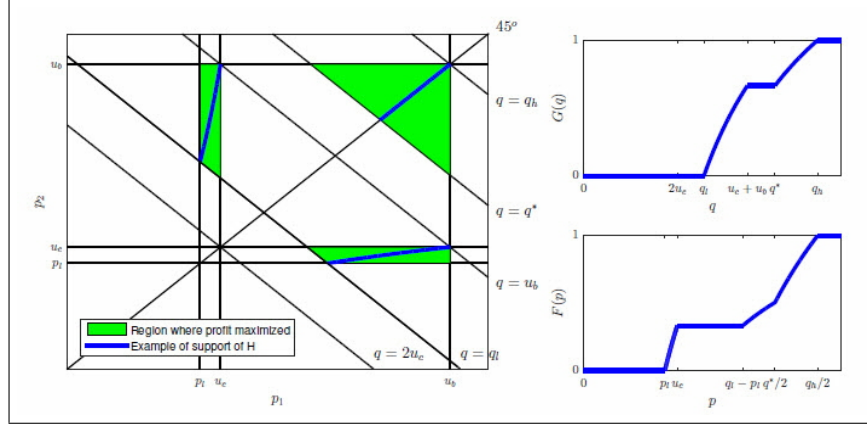
$$p_l = \frac{\alpha + 2(1 - \alpha)[1 - G(q^*)/2]}{2 - \alpha} u_c. \quad (8)$$

This completes the characterization of the equilibrium. In this equilibrium, there is a group of sellers who sets a basket price of $q \in [q^*, q_h]$ and the prices p_1 and p_2 in between u_c and u_b . These sellers trade (with some probability) the basket of goods to buyers of type b and never trade with buyers of type c . There is also a group of sellers who set a basket price of $q \in [q_l, u_b + u_c]$. Half of these sellers set p_1 below u_c and p_2 between u_c and u_b . These sellers trade (with some probability) the whole basket of goods to buyers of type b and good 1 to buyers of type c . The other half of the sellers sets p_2 below u_c and p_1 between u_c and u_b . These sellers trade (with some probability) the whole basket of goods to buyers of type b and good 2 to buyers of type c . There are no sellers who set a basket price of q in the interval $(u_b + u_c, q^*)$.

The distribution of basket prices $G(q)$ is given by equation (1) for $q \in [q^*, q_h]$ and by equation (5) for $q \in [q_l, u_b + u_c]$. The distribution $G(q)$ is such that the seller's profit from trading the basket of goods to buyers of type b is equal to S^* for all $q \in [q^*, q_h]$, and it is equal to $S^* - \mu_c[\alpha + 2(1 - \alpha)(1 - F(u_c))]u_c$ for all $q \in [q_l, u_b + u_c]$. The distribution $G(q)$ has a gap between $u_b + u_c$ and q^* . The gap exists because a seller with a basket price of $u_b + u_c$ trades with both buyers of type b and buyers of type c , while a seller with a basket price greater than $u_b + u_c$ only trades with buyers of type b . Therefore, a seller strictly prefers setting a basket price of $u_b + u_c$ rather than setting any basket price just above $u_b + u_c$. The distribution of prices for an individual good $F(p)$ is given by equation (7) for $p \in [p_l, u_c]$. The distribution $F(p)$ is such that the seller's profit from trading the cheaper good to buyers of type c is equal to $S^* - \mu_b(2 - \alpha)q_l$ for all $p \in [p_l, u_c]$. The distribution $F(p)$ is not uniquely pinned down for $p \in (u_c, u_b]$. Intuitively, this is the case because a seller who charges a price of $p > u_c$ for one good only trades that good to buyers of type b together with the other good.

The distribution of price vectors H is not uniquely pinned down. For sellers with a basket price $q \in [q^*, q_h]$, there are several distributions H that generate the marginal distribution of basket prices $G(q)$ in equation (1) and thus are consistent with equilibrium. For example, as discussed in KMRT, there is an equilibrium in which, for all $q \in$

Figure 2 Equilibrium with Relative Price Dispersion (from KMRT)



Notes: This figure shows the possible range of the support of the joint distribution $H(p_1, p_2)$, the shape of the cumulative distributions $G(q)$, and an example of the shape of the cumulative distribution $F(p)$ in the discrimination equilibrium.

$[q^*, q_h]$, there are $G'(q)$ sellers with a basket price of q , and each of them posts the prices $(q/2, q/2)$. For sellers with a basket price $q \in [q_l, u_b + u_c]$, there are again several distributions H that generate the marginal distribution of basket prices $G(q)$ in equation (5) and the marginal distribution of individual good prices $F(p)$ in equation (7) that are consistent with equilibrium. For example, there is an equilibrium in which, for all $p \in [p_l, u_c]$, $2F'(p)$ sellers have a basket price of $\phi(p)$, $F'(p)$ sellers post the prices $(p, \phi(p) - p)$, and $F'(p)$ sellers post the prices $(\phi(p) - p, p)$, where

$$\phi(p) = \frac{[\alpha + 2(1 - \alpha)(1 - G(q^*))](u_b + u_c)}{[\alpha + 2(1 - \alpha)(1 - G(q^*))] + 2[\alpha + 2(1 - \alpha)(1 - G(q^*)/2)](u_c - p)/2}.$$

A graphical representation is provided in Figure 2.

To conclude the analysis, it is necessary to provide necessary and sufficient conditions for the existence of this equilibrium. The equilibrium exists if and only if

$$\frac{\mu_c}{\mu_b} > \frac{3\alpha - 2}{(2 - \alpha)u_c/u_b} - 1, \quad \frac{\mu_c}{\mu_b} \leq \frac{\alpha - (2 - \alpha)u_c/u_b}{1 + (2 - \alpha)u_c/u_b} \frac{1 + u_c/u_b}{u_c/u_b}. \quad (9)$$

The first condition guarantees that some sellers find it optimal to post basket prices below $u_b + u_c$. The condition is satisfied if: (i) the market is sufficiently competitive, in the sense that the fraction α of buyers who are in contact with only one seller is smaller than $2/3$; or (ii) the relative number of type c buyers, μ_c/μ_b , and/or the relative willingness to pay of type c buyers, u_c/u_b , is large enough. The second condition guarantees that no seller finds it optimal to post prices below $2u_c$. The condition is satisfied if: (i) the market is not too competitive, in the sense that the fraction α of buyers who are in contact with only one seller is greater than $2(u_c/u_b)/(1+2(u_c/u_b))$; or (ii) the relative number of type c buyers, μ_c/μ_b , and/or the relative willingness to pay of type c buyers, u_c/u_b , is low enough.

The next proposition summarizes the equilibrium discussed in this section.

Proposition 1 *The equilibrium exists if the conditions in equation (9) are satisfied. In the equilibrium, the bundle price distribution G is continuous on the support $[q_l, u_b + u_c][[q^*, q_h]$, where q^* is given by equation (4), q_l is given by equation (6), and $q_h = 2u_b$. For $q \in [q^*, q_h]$ we have that G is given by equation (1), while it is given by equation (5) for $q \in [q_l, u_b + u_c]$. The distribution of individual prices F is continuous on the interval $[p_l, u_c]$, where p_l is presented in equation (8), and it is given by equation (7).*

Discussion

The equilibrium features price dispersion across sellers, in the sense that some sellers are on average expensive, while some sellers are on average cheap. This property of equilibrium follows immediately from the fact that the distribution of basket prices is nondegenerate. A discrimination equilibrium always features relative price dispersion, in the sense that there is variation across sellers in the price of a particular good at a particular seller relative to the average price charged by that seller. This property of equilibrium follows immediately from the fact that half of the sellers with a basket price $q \in [q_l, u_b + u_c]$ have a relative price for good 1 that is strictly greater than 1, while the other half of the sellers with a basket price $q \in [q_l, u_b + u_c]$ have a relative price for good 1 that is strictly smaller than 1.

Why does relative price dispersion emerge in equilibrium? Competition between sellers drives part of the distribution of basket prices to the region where q is between $2u_c$ and $u_b + u_c$. A seller with a basket price between $2u_c$ and $u_b + u_c$ never finds it optimal to post the same price for both goods. Instead, the seller finds it optimal to set the price

of one good below and the price of the other good above the willingness to pay of type c buyers. That is, a seller with a basket price q between $2u_c$ and $u_b + u_c$ finds it optimal to follow an asymmetric pricing strategy for the two goods. However, if some sellers post a higher price for good 1 than for good 2, other sellers must post a higher price for good 2 than for good 1, or else there would be some unexploited profit opportunities. That is, the distribution of prices for the two goods must be symmetric across sellers with a basket price q between $2u_c$ and $u_b + u_c$. The asymmetric pricing strategy followed by each individual seller combined with the symmetry of the price distribution across sellers implies relative price dispersion.

Sellers follow an asymmetric pricing strategy to discriminate between the two types of buyers. The difference in the willingness to pay of type b and type c buyers gives sellers a desire to price discriminate. The difference in the ability of type b buyers and type c buyers to purchase different items in different locations gives sellers the opportunity to price discriminate. In fact, by pricing the two goods asymmetrically, a seller can charge a high average price to the high-valuation buyers who need to purchase all the items together (the buyers of type b) and charge a low price for one good to the low-valuation buyers who can purchase different items at different locations (the buyers of type c).

3. CONCLUDING REMARKS

In this paper, I reviewed the work by Kaplan et al. (2016). The paper studies price dispersion both empirically and theoretically in setting where firms sell (and price) multiple goods. Empirically, the paper finds that an important fraction of price dispersion for identical goods is due to relative price dispersion. That is, due to the fact that stores with the same overall price level sell individual goods in a persistently different way. The paper then describes a theory that can rationalize its empirical findings, relying on stores that sell multiple goods trying to price discriminate heterogeneous customers.

Although the equilibrium is unique, the fact that it is displayed as a discrimination equilibrium depends on the parameters of the model, as described in equation (9). In fact, when the first condition is not satisfied (for example, when the fraction of cool shoppers is low), the equilibrium is such that stores only sell baskets of goods to busy shoppers and, as previously discussed, individual prices would not be pinned down in equilibrium, and thus relative price dispersion would not be a robust prediction of the model. When the first condition is satisfied and the second condition is not satisfied (for example, when the fraction of cool shoppers is moderately high), at least some stores are

willing to sell both goods to cool shoppers, and thus these stores act as unbundled. Still, relative price dispersion survives here as some stores still price discriminate. Finally, when the fraction of cool shoppers is big enough, the equilibrium becomes completely unbundled, with every store attempting to sell both goods to both cool and busy shoppers.

It is interesting to contrast the type of price discrimination advanced in Kaplan et al. (2016) with intertemporal price discrimination (see, e.g., Conlisk et al. [1984] and Sobel [1984] or, in a search-theoretic context, Albrecht et al. [2013] and Menzio and Trachter [2015b]). The key to intertemporal price discrimination is a negative correlation between a buyer's valuation and his ability to intertemporally substitute purchases. A seller can exploit this negative correlation by having occasional sales. The low-valuation buyers, who are better able to substitute purchases intertemporally, will take advantage of the sales and will end up paying low prices. The high-valuation buyers, who are unable to substitute purchases intertemporally, will not take advantage of the sales and will end up paying high prices. In contrast, this theory of price discrimination is based on a negative correlation between a buyer's valuation and his ability to shop in multiple stores. Moreover, while intertemporal price discrimination takes the form of time variation in the price of the same good, this theory of price discrimination takes the form of variation in the price of different goods relative to the average store price.

REFERENCES

- Aguirregabiria, Victor. 1999. "The Dynamics of Markups and Inventories in Retailing Firms." *Review of Economic Studies* 66 (April): 275–308.
- Albrecht, James, Fabien Postel-Vinay, and Susan Vroman. 2013. "An Equilibrium Search Model Of Synchronized Sales." *International Economic Review* 54 (May): 473–93.
- Burdett, Kenneth, and Kenneth L. Judd. 1983. "Equilibrium Price Dispersion." *Econometrica* 51 (July): 955–69.
- Campbell, Jeffrey R., and Benjamin Eden. 2014. "Rigid Prices: Evidence From U.S. Scanner Data." *International Economic Review* 55 (May): 423–42.
- Conlisk, John, Eitan Gerstner, and Joel Sobel. 1984. "Cyclic Pricing by a Durable Goods Monopolist." *Quarterly Journal of Economics* 99 (August): 489–505.
- Ellison, Sara Fisher, Christopher M. Snyder, and Hongkai Zhang. 2015. "Costs of Managerial Attention and Activity as a Source of Sticky Prices: Structural Estimates from an Online Market." Mimeo, Massachusetts Institute of Technology.
- Gorodnichenko, Yuriy, Viacheslav Sheremirov, and Oleksandr Talavera. 2015. "Price Setting in Online Markets: Does IT Click?" Federal Reserve Bank of Boston Working Paper 15-1 (January).
- Kaplan, Greg, and Guido Menzio. 2015. "The Morphology Of Price Dispersion." *International Economic Review* 56 (November): 1165–1206.
- Kaplan, Greg, Guido Mezio, Leena Rudanko, and Nicholas Trachter. 2016. "Relative Price Dispersion: Evidence and Theory." Federal Reserve Bank of Richmond Working Paper 16-02 (January).
- Lewis, Matthew. 2008. "Price Dispersion and Competition with Differentiated Sellers." *Journal of Industrial Economics* 56 (September): 654–78.
- Menzio, Guido, and Nicholas Trachter. 2015a. "Equilibrium Price Dispersion Across and Within Stores." National Bureau of Economic Research Working Paper 21493 (August).
- Menzio, Guido, and Nicholas Trachter. 2015b. "Equilibrium Price Dispersion with Sequential Search." *Journal of Economic Theory* 160 (December): 188–215.

Sobel, Joel. 1984. "The Timing of Sales." *Review of Economic Studies* 51 (July): 353–68.

Monetary Incentives and Mortgage Renegotiation Outcomes

Nika Lazaryan and Urvi Neelakantan

1. INTRODUCTION

The U.S. foreclosure crisis began in 2006, when over 700,000 properties received foreclosure filings (RealtyTrac Staff 2014). The number of filings increased every year until 2010, at which time they peaked at nearly 2.9 million. The inventory of mortgage foreclosures as a share of outstanding mortgages increased from around 1 percent in 2000 to 4.6 percent in 2010.¹ The historically unprecedented numbers prompted the U.S. government to introduce several programs to reduce the number of foreclosures.² Prominent among these programs was the Home Affordable Modification Program (HAMP), which was introduced in 2009. Its goal was to help homeowners avoid foreclosure by encouraging servicers to work with homeowners to modify the terms of their mortgage. HAMP offered servicers \$1,000 for each modification

■ The idea for this paper germinated from a collaboration in 2010–12 with Community Development staff members at the Federal Reserve Bank of Richmond, whom we acknowledge with gratitude. This is a companion paper to Neelakantan et al. (2012), coauthored with Shannon McKay and Kim Zeuli, which assesses the predictions of the theoretical model presented in the current piece using survey data on homeowners who sought mortgage assistance. We thank Andreas Hornstein, Erica Paulos, Ned Prescott, Pierre-Daniel Sarte, Russell Wong, and seminar participants at the Federal Reserve Bank of Richmond for helpful comments. 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. Nika Lazaryan: Federal Reserve Bank of Richmond, P.O. Box 27622, Richmond, VA 23261, Nika.Lazaryan@rich.frb.org, Ph:804-697-5475. Urvi Neelakantan: Federal Reserve Bank of Richmond, P.O. Box 27622, Richmond, VA 23261, Urvi.Neelakantan@rich.frb.org, Ph:804-697-8146.

¹ Data from Mortgage Bankers Association via Haver Analytics.

² See Gerardi and Li (2010) for a discussion of these programs. Prior to 2007, there were no federal programs addressing mortgage default (Hembre 2014).

completed under the program (Making Home Affordable Program 2010). Additional incentives were offered to homeowners and servicers for up to three years for loans that remained in good standing.³ Compared to regular servicing fees of 20 to 50 basis points of the outstanding loan balance, these incentives were quite sizable.⁴

The goal of this paper is to examine the effect of incentives on mortgage renegotiation or modification (the terms are used interchangeably) outcomes. Specifically, we are interested in whether incentives offered to homeowners and servicers can indeed reduce foreclosures.⁵ To address this question, we use a simple model of renegotiation between the homeowner and lender. The model is a sequential-move game in which the homeowner moves first and decides whether to seek renegotiation. Next, the lender decides whether to modify the terms of the mortgage. The homeowner then decides whether to default. Homeowners who default are foreclosed upon. We compare the predictions of the model with no incentives to predictions of the model in which incentives are introduced.

Results show that, in the absence of incentives, lenders would renegotiate only with the subset of homeowners who would neither i) redefault despite receiving modified terms nor ii) self-cure without modified terms. (The ideas of “self-cure” and “redefault” are formalized in the model.) The renegotiation enables this subset of homeowners to avoid foreclosure. Once incentives are introduced, the subset of homeowners who receive renegotiated terms and avoid foreclosure is larger than the subset in the model without incentives. However, if incentive payments to the lender are sufficiently high, we find that lenders may also renegotiate with homeowners they know will subsequently redefault. To summarize, we find that incentives can indeed reduce the number of foreclosures, but there are scenarios in which some of the incentive payments are channeled to renegotiations in which foreclosure is still the final outcome. Note that these are descriptive results; assessing the costs and benefits or the welfare implications of such outcomes, or of the particulars of the HAMP program, is beyond the scope of this paper.⁶

³ The ongoing “pay-for-success” incentives included up to \$1,000 in yearly payments for three years after the modification for the borrowers who were current on their mortgage payments.

⁴ Regular servicing fees on a mortgage with a \$200,000 balance are between \$400 and \$1,000 per year (Agarwal et al. 2012).

⁵ We use the term “servicer” and “lender” interchangeably in the remainder of the paper, because the distinction is not relevant for our model.

⁶ For an assessment of the net benefits and the effectiveness of the HAMP program in particular, see Hembre (2014) and Scharlemann and Shore (2016).

2. RELATED LITERATURE

We rely on the literature to motivate key assumptions in our model. Our first assumption is that homeowners have negative equity in their home, i.e., their mortgage balance exceeds the price of their house. When the borrower has positive equity in the property, it may not be optimal for them to default, especially if they can sell the property, pay off the mortgage, and keep or use the difference (Foote et al. 2008, 2010). There is strong empirical evidence that borrower defaults happen in conjunction with negative equity (Deng et al. 2000; Danis and Pennington-Cross 2008; Gerardi et al. 2008; Campbell and Cocco 2015; Goodman et al. 2010; Ghent and Kudlyak 2011). The foreclosure crisis was characterized by falling house prices, which increased the number of borrowers with negative equity in their homes. Campbell et al. (2011) argue that foreclosures exacerbated the house price decline by negatively affecting the prices of neighboring houses, further increasing the number of borrowers faced with negative equity. However, negative equity alone does not always imply that the borrower should choose to default (Deng et al. 2000; Foote et al. 2008, 2010). We allow for this by making default costly — in principle, the negative impact on the borrower’s credit history, potential relocation costs, and other monetary and non-monetary costs can deter even those borrowers with negative equity from defaulting.⁷ We allow the cost of default to vary across borrowers in our model. As will become clear in the model section, this leads to borrowers of three broad types: those who self-cure (i.e., become current on their loan without receiving modified terms), those who redefault (i.e., default again after receiving a mortgage modification), and those in between (i.e., those who default without modified terms but remain current after a modification).

The fact that lenders have to face borrowers of different types has been cited as a reason for lenders’ reluctance to renegotiate mortgages (White 2009a, 2009b; Adelino et al. 2013; Ghent 2011). Since renegotiation does not guarantee that the borrower will not default again in the future, the lender would not want to renegotiate mortgage terms with borrowers who would subsequently redefault on the loan. If they did, the lender would not only incur the losses associated with foreclosure, but also lose additional funds associated with the cost of renegotiations. Conversely, the lender would also not want to renegotiate with borrowers who could self-cure, since the modified terms would lead to

⁷ The literature suggests that default is the result of a “double trigger”—negative home equity in conjunction with an adverse shock affecting the borrower’s ability to make payments (see, for example, Gerardi et al. 2013; Elul et al. 2010). Our simple model abstracts from such adverse shocks.

a loss of revenue for the lender without any offsetting benefits. In the cost-benefit analysis of Ambrose and Capone (1996), when either the probability of self-cure or redefault is sufficiently high, it is no longer optimal for the lender to consider loan renegotiation as an option. In fact, recent empirical evidence shows that these two categories comprise a sizeable portion of the borrowers.⁸ Thus, as pointed out by Adelino et al. (2013), in the presence of uncertainty about borrower types, lenders could prefer to foreclose. The goal of our analysis is to assess whether and how incentive payments change this calculus.

We model the renegotiation between the homeowner and lender as a sequential move game, which is consistent with previous literature (Adelino et al. 2013; Wang et al. 2002). A key difference is that, while prior works highlight the role of information asymmetry as a barrier to successful renegotiations, we aim to uncover issues that might arise even with full information in the presence of incentives.⁹ Our contribution is thus to assess the effectiveness of incentives absent any other barriers to renegotiation. We also provide a simple theoretical underpinning for empirical observations about programs such as HAMP. For example, certain parameterization of our model can explain why lenders renegotiate only a small fraction of delinquent loans, as pointed out by Adelino et al. (2013).¹⁰ In the presence of incentives, our model predicts that the subset of homeowners who receive a modification and avoid foreclosure is larger. This is consistent with Agarwal et al. (2012) and Scharlemann and Shore (2016), who find that HAMP led to a modest reduction in foreclosures. Papers that focus on recent modification programs find that these programs attract homeowners who might otherwise self-cure (see, for example, Mayer et al. 2014), which is also a result that our model delivers. In addition, we characterize parameters of the model under which lenders renegotiate with homeowners who subsequently redefault.

⁸ Adelino et al. (2013) look at the sample of mortgages from 2005–08 and find that more than 30 percent of seriously delinquent borrowers end up becoming current on their mortgages without receiving any mortgage modification. On the other hand, around 20 to 50 percent of the borrowers default after receiving loan modification.

⁹ In the presence of information asymmetry, lenders can choose to incur screening costs to distinguish between borrower types. Wang et al. (2002) show that the optimal policy of the lender in this case is to either: 1) screen through enough applications so that borrowers who could self-cure are discouraged from seeking assistance, or 2) to randomly reject requests for mortgage modification, at a rate that depends on liquidation cost and magnitude of default, among other factors.

¹⁰ Data on the HAMP program suggests that this might be the case for HAMP as well: as of February 2014, servicers had processed over 7.7 million applications but have approved less than one-third of them (Making Home Affordable Program 2010).

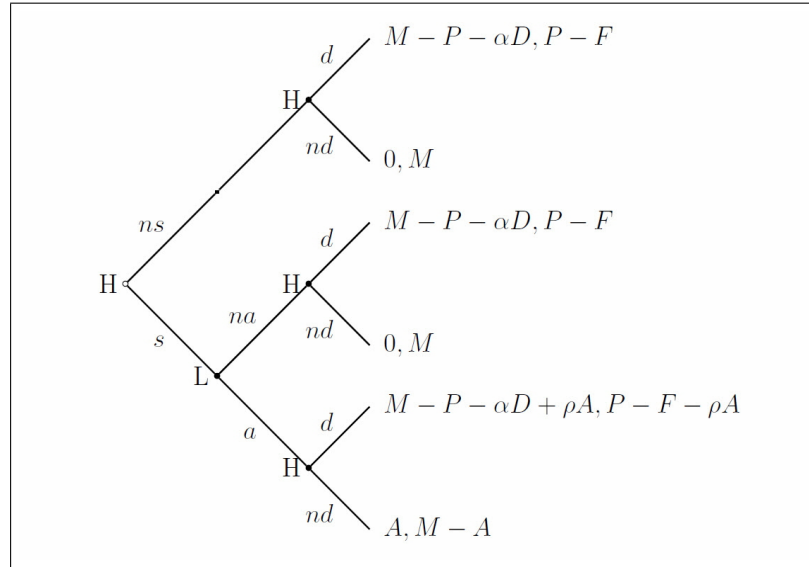
3. THE MODEL

In the model of strategic interaction, the players are a single lender and a continuum of homeowners of type α , where α is uniformly distributed on the interval $[0, 1]$. Let M denote the mortgage balance and P the market price of the home. It is assumed that $M - P > 0$, based on the literature that finds that negative equity is a trigger for default, e.g., Foote et al. (2008).

Figure 1 illustrates the payoffs of the possible outcomes of the interaction between the lender (L) and an individual homeowner (H). The homeowner moves first and decides whether to seek renegotiation (denoted by action s) or not seek renegotiation (ns). If he does not seek renegotiation and does not default on his mortgage (denoted by action nd), there is no change to his present situation and his payoff is 0. If he defaults (denoted by action d), he is foreclosed upon and his payoff is $M - P - \alpha D$. This is because he loses the house, whose market value is P , but no longer has to pay the mortgage M . For the homeowner of type α , the cost of defaulting is αD . This expression reflects the assumption that homeowners differ in their cost of mortgage default. If the homeowner does not seek renegotiation and does not default, the lender receives the mortgage amount M as per the original contract. If he defaults and is foreclosed upon, the lender takes possession of the house. Her payoff is the market value P of the house less the cost associated with foreclosing on it, F .

Once the homeowner decides to seek renegotiation, the lender has to decide whether or not to agree. If the lender does not agree to renegotiate (na), the homeowner's payoffs are the same as in the case where he chose not to seek renegotiation. Thus the payoff to the homeowner of seeking but not receiving a modification and then not defaulting is 0, while the payoff from defaulting is $M - P - \alpha D$. There is no change to the lender's payoff either; she receives M if the homeowner does not default and $P - F$ if he does.

If the lender agrees, denoted by action a , the modification leads to the homeowner being paid an amount A . If the homeowner does not default, his payoff is A . In this case, the lender receives $M - A$. If the homeowner receives A and still defaults, his payoff is $M - P - \alpha D + \rho A$. Since there is no time dimension in the model, $\rho \in (0, 1)$ loosely captures what might occur during the modification process. Consider an example in which a homeowner receives a lower interest rate. We can think of the total amount A as the difference between the original payments and the new, lower payments under the new interest rate over the full length of the loan term. However, if the homeowner defaults and is foreclosed upon after making a few of the new payments, he

Figure 1 Homeowner and Lender Payoffs

receives in effect only a fraction of the amount, i.e., ρA . In this case, the lender's payoff is $P - F - \rho A$.

Model with No Incentives

We first assume that there is no government program in place. In other words, renegotiations between the lender and homeowner are purely bilateral with no externally funded incentives.

In principle, it is possible for the lender to choose both whether or not to renegotiate and how much to offer the homeowner. However, to avoid the complexities associated with a continuum of strategies, we assume for now that the lender has only two choices — not renegotiate (*na*) or agree to renegotiate and offer a specific amount $A = M - P$.¹¹ The payoffs under this specific assumption are shown in Figure 2.

In solving this game backward, we observe that homeowners can be grouped into types. Some homeowners would not default at any of

¹¹ This assumption follows Wang et al. (2002). Letting $A = M - P$ assumes in effect that the lender eliminates the homeowner's negative equity. Such a policy has actually been proposed and is critiqued in Gerardi and Willen (2009).

the terminal nodes. For these homeowners, $\alpha \in [\bar{\alpha}, 1]$, where

$$\bar{\alpha} = \frac{M - P}{D}. \quad (1)$$

Also observe that there are homeowners who would get a higher payoff from defaulting even when offered A . For these homeowners, $\alpha \in [0, \underline{\alpha}]$, where

$$\underline{\alpha} = \frac{\rho(M - P)}{D}. \quad (2)$$

We assume that $0 < \underline{\alpha} < \bar{\alpha} < 1$. In other words, homeowners can be grouped into three categories: (i) those with $\alpha \in [0, \underline{\alpha}]$ who would default even if they received a modification, (ii) those with $\alpha \in [\bar{\alpha}, 1]$ who would not default even if they received no modification, and (iii) those with $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ who would default if they received no modification but not if they received a modification.

In the absence of any renegotiation between the lender and homeowners, all homeowners with $\alpha \in [0, \bar{\alpha})$ would default on their mortgages and be foreclosed upon while all homeowners with $\alpha \in [\bar{\alpha}, 1]$ would not. The lender's payoff in this case would be

$$\bar{\alpha}(P - F) + (1 - \bar{\alpha})M. \quad (3)$$

We now formally describe the solution to the model by characterizing the subgame perfect Nash equilibrium. This requires specifying the strategy profile that includes strategies of every player. Since there is a continuum of homeowners, we describe strategy profiles over intervals within $[0, 1]$.

Proposition 1 *Assume full information (the homeowners' type and the lenders' actions are observable). Let $\underline{\alpha} = \frac{\rho(M-P)}{D}$ and $\bar{\alpha} = \frac{M-P}{D}$. Then the strategy profile¹²*

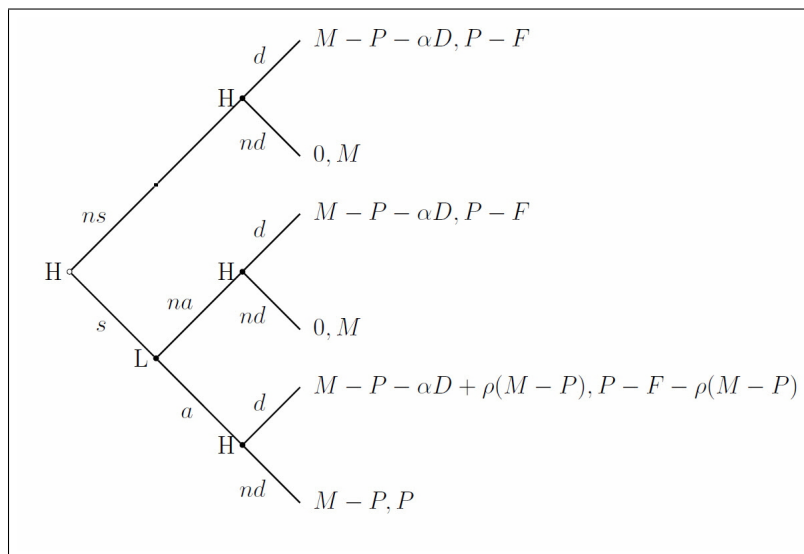
$$\begin{aligned} &\{(s \text{ Always choose } d), na\} \quad \forall \alpha \in [0, \underline{\alpha}] \\ &\{(s \text{ } nd|A = M - P \text{ } d|otherwise), a\} \quad \forall \alpha \in [\underline{\alpha}, \bar{\alpha}] \\ &\{(s \text{ Always choose } nd), na\} \quad \forall \alpha \in [\bar{\alpha}, 1] \end{aligned}$$

*is a subgame perfect Nash equilibrium of the game in Figure 2.*¹³

Proof. See Appendix. ■

¹² The strategy profile is of the form {(Homeowner's strategy at initial node Homeowner's conditional strategy at terminal nodes), Lender's strategy}.

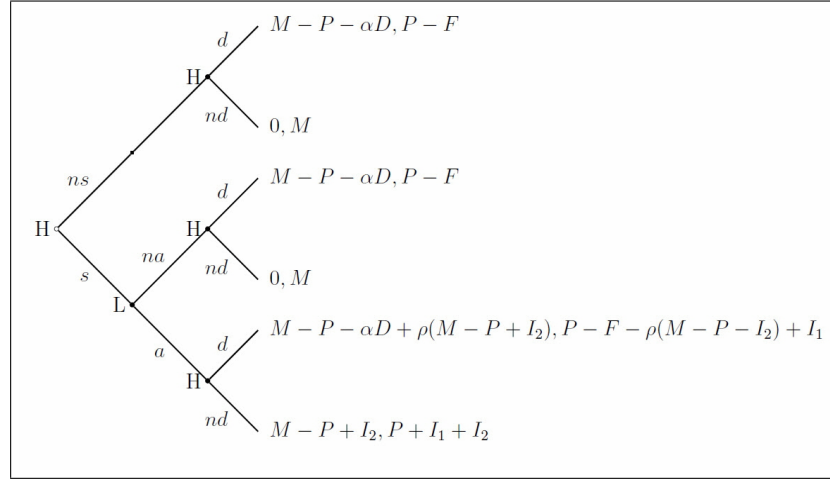
¹³ Note that the subgame perfect Nash equilibrium is not unique. To be specific, strategy profiles in which homeowners with $\alpha \in [0, \underline{\alpha}]$ and $\alpha \in [\bar{\alpha}, 1]$ always chose action ns , or randomize between s and ns , would also be subgame perfect Nash equilibria because the payoffs from the two are the same.

Figure 2 Homeowner and Lender Payoffs with $A=M-P$ 

The preceding result shows that there is an equilibrium in which all types of homeowners choose to seek renegotiation. This illustrates the point that Adelino et al. (2013) make: renegotiation exposes the lender to homeowners who would self-cure (those with $\alpha \in [\bar{\alpha}, 1]$ in our model) or redefault (those with $\alpha \in [0, \underline{\alpha})$). The lender does not renegotiate with homeowners of type $\alpha \in [0, \underline{\alpha})$ because they would default even if they received a modification. As a result, the lender's payoff from renegotiating, $P - F - \rho A$, would be strictly less than her payoff from not doing so, $P - F$. The lender also does not renegotiate with homeowners of type $\alpha \in [\bar{\alpha}, 1]$ because her payoff from not modifying the terms, M , is strictly higher than her payoff from modifying the terms, $M - A$. In this equilibrium, the only homeowners whose mortgage terms are modified are of type $\alpha \in [\underline{\alpha}, \bar{\alpha})$. These are homeowners who would have gone through foreclosure in the absence of the modification but avoid foreclosure because they receive it.

It can be shown that the payoff to the lender from the above solution exceeds the payoff from the solution with no renegotiation as described by equation (3).

Certain parameterizations of the model can yield results consistent with empirical observations. For example, Adelino et al. (2013) point out that lenders renegotiate only a small fraction of delinquent loans.

Figure 3 Homeowner and Lender Payoffs with Incentives

Our model can obtain a qualitatively similar result if the interval $[\underline{\alpha}, \bar{\alpha})$ is small, that is, if the number of homeowners who would successfully avoid foreclosure with a modification is small relative to the number who would redefault or self-cure.

Model with Incentives

We now solve the model in the presence of a government program that gives incentives to homeowners and lenders. We are particularly interested in comparing the solutions from this model to the model without the program to see whether the former is more effective in terms of preventing foreclosure.

The model of homeowner and lender renegotiation in the presence of incentives is shown in Figure 3. Our modeling of incentives is motivated by HAMP rules that were in place in 2010. Specifically, the program offered incentive compensation of \$1,000 to servicers for each permanent modification completed (Making Home Affordable Program 2010). In addition, it offered up to \$1,000 each to the homeowner and servicer for every year that the loan remained in good standing (or \$83.33 monthly), for a maximum of three years. We introduce this incentive compensation structure into our model as follows. The lender receives I_1 for offering a modification, regardless of whether or not the

homeowner subsequently defaults. If the homeowner does not default and thereby avoids foreclosure, the lender receives an additional I_2 as “pay-for-success.” As before, we use ρ to capture what might happen during the modification period. In particular, if the homeowner remains current for a few periods after the renegotiation, both the homeowner and the lender would receive partial pay-for-success payments ρI_2 .

To compare the solution from this model to the model with no incentives, assume that all other variables are the same as before. We first show that an equilibrium exists in which a larger fraction of homeowners receives modifications and avoids foreclosure. The incentives thus have the effect of preventing some foreclosures that would have occurred in the absence of the program. The following result characterizes the equilibrium.

Proposition 2 *Assume full information. Let $\underline{\alpha}' = \frac{\rho(M-P)-(1-\rho)I_2}{D}$ and $\bar{\alpha} = \frac{M-P}{D}$. Assume that $\rho(M-P) \geq (1-\rho)I_2$, that $\rho(M-P-I_2) > I_1$, and that $I_1 + I_2 < M - P$. Then the strategy profile*

$$\begin{aligned} &\{(s \text{ Always choose } d), na\} \quad \forall \alpha \in [0, \underline{\alpha}'] \\ &\{(s \text{ } nd|A = M - P \text{ } d|\text{otherwise}), a\} \quad \forall \alpha \in [\underline{\alpha}', \bar{\alpha}] \\ &\{(s \text{ Always choose } nd), na\} \quad \forall \alpha \in [\bar{\alpha}, 1] \end{aligned}$$

*is a subgame perfect Nash equilibrium of the game in Figure 3.*¹⁴

Proof. See Appendix. ■

Comparing Proposition 2 to Proposition 1, we see that the results are qualitatively similar. All homeowners seek renegotiation, but the lender offers it only to the subset of homeowners who can successfully avoid foreclosure as a result. The key difference is that the subset of homeowners who receive a modification and avoid foreclosure is larger in this case. This follows from the fact that $\underline{\alpha}' < \underline{\alpha}$. Intuitively, the homeowners’ payoff from receiving a modification and not defaulting is increased by the incentive payment I_2 , which makes this option attractive to a larger fraction of homeowners.

The next result shows that, under different assumptions about the incentive structure, lenders may be induced to also renegotiate with homeowners of type $\alpha \in [0, \underline{\alpha}']$, and that these homeowners will subsequently default.

¹⁴ For the same reasons as described for Proposition 1, the equilibrium is not unique.

Proposition 3 *Assume full information. Let $\underline{\alpha}' = \frac{\rho(M-P)-(1-\rho)I_2}{D}$ and $\bar{\alpha} = \frac{M-P}{D}$. Assume that $\rho(M-P) \geq (1-\rho)I_2$, that $I_1 \geq M-P-I_2$, and that $I_1 + I_2 < M - P$. Then the strategy profile*

$$\begin{aligned} &\{(s \text{ Always choose } d), a\} \quad \forall \alpha \in [0, \underline{\alpha}'] \\ &\{(s \text{ } nd|A = M - P \text{ } d|otherwise), a\} \quad \forall \alpha \in [\underline{\alpha}', \bar{\alpha}] \\ &\{(ns \text{ Always choose } nd), na\} \quad \forall \alpha \in [\bar{\alpha}, 1] \end{aligned}$$

is a subgame perfect Nash equilibrium of the game in Figure 3.

Proof. See Appendix. ■

As in Proposition 2, a larger fraction of homeowners receives modifications and avoids foreclosure compared to the no-incentive case. The key difference between this result and Proposition 2 is that the lender now also renegotiates with all homeowners of type $\alpha \in [0, \bar{\alpha}]$. Homeowners of this type subsequently default and are foreclosed upon. The reason for the difference in the two results is the incentive structure. In particular, the incentive payment given to the lender simply for renegotiating, I_1 , is higher than in the previous case and also higher than the pay-for-success incentive I_2 (this follows from the assumptions in Proposition 3). This makes it worthwhile for the lender to renegotiate even with those homeowners who default.¹⁵

Proposition 3 highlights the fact that the parameters of the incentive structure can make the program less effective, in the sense of allocating some incentives to renegotiations that still result in foreclosure. This can happen, for example, if the pay-for-success payment, I_2 , is not much higher than the incentive to participate, I_1 , and if the homeowner redefaults fairly quickly, i.e., if ρ is also low.

Finally, observe that it is possible in theory but unlikely in practice to have incentives large enough to induce lenders to renegotiate with homeowners who would otherwise self-cure. This can be seen if the proof of Proposition 2 was reworked under the assumption that $I_1 + I_2 \geq M - P$. This is an unlikely assumption in practice because it requires that the incentive payments exceed the modification amount that the lender offers.

To summarize, our models show that in the absence of incentives, the lender renegotiates the mortgage terms of a subset of homeowners who avoid foreclosure as a result. In the presence of incentives, the lender renegotiates with a larger subset of homeowners who avoid foreclosure as a result. However, under certain assumptions about the

¹⁵ Mayer et al. (2009) propose an incentive fee structure that would avoid this scenario by rewarding servicers only for successful modifications.

incentive structure, the lender may also renegotiate with homeowners who subsequently default and are foreclosed upon.

Mortgage Modifications and Success Rates

Mortgage modifications are often evaluated by comparing “success rates” — defined as the fraction of homeowners who avoid foreclosure — across homeowners who do and do not receive modifications. Our models show that this comparison is not necessarily informative about the effectiveness of mortgage modifications. This is because success rates among those who do not receive modifications may be high if this group includes a large proportion of homeowners who self-cure. The solutions described by Proposition 1 and Proposition 2 illustrate this. In those solutions, the success rate conditional on not receiving a modification is $\frac{1-\bar{\alpha}}{1-\bar{\alpha}+\underline{\alpha}}$. This number can be close to 1 if the interval $[\bar{\alpha}, 1]$ is large relative to the interval $[0, \underline{\alpha}]$. Recent research suggests that this is indeed the case. For example, Mayer et al. (2014) find that borrowers who became delinquent following a program announcement to help seriously delinquent borrowers were “those who appear to have been least likely to default otherwise.”¹⁶ As a result, cure rates or success rates can end up being high among those who do apply but do not receive modifications. The conclusion is that success rate comparisons should be interpreted with caution when judging the effectiveness of mortgage modification programs.

4. CONCLUSION

The model in this paper provides a simple framework to analyze mortgage renegotiation between homeowner and lender. The results allow for a comparison of outcomes in the absence of incentives to outcomes in the presence of externally funded incentives to homeowners and lenders. In the absence of incentives, lenders renegotiate only with those homeowners who would successfully avoid foreclosure upon receiving a modification but would default without it. In other words, lenders do not renegotiate with homeowners who would self-cure without a modification or with homeowners who would default despite receiving it. The share of homeowners who receive modifications and avoid foreclosure is larger in the presence of incentives, and in some cases incentives might also induce lenders to renegotiate with homeowners who subsequently default. It is beyond the scope of this paper

¹⁶ Andersson et al. (2013) also suggest that HAMP may have made default on mortgage debt more attractive.

to determine whether the benefit exceeds the cost of providing such incentives or the overall impact of such programs on foreclosure prevention.

An important caveat is that this paper abstracts from information asymmetry between the lender and homeowner. We think that is a reasonable abstraction that enables us to focus on considerations even in the presence of full information. As Agarwal et al. (2012) describe, HAMP, for example, had extensive screening criteria, including trial periods, that likely enabled lenders to learn a lot about the homeowners. However, to the extent that asymmetric information is an issue, it may overstate how much lenders are able to target the “right” homeowners. Nonetheless, the point we illustrate is that even if lenders are able to target the right homeowners, externally funded incentives may lead them to also renegotiate with homeowners who cannot be protected from foreclosure.

APPENDIX

Proof of Proposition 1

Proof. We show that the strategy profile is a subgame perfect Nash equilibrium by solving the game in Figure 2 by backward induction.

For homeowners of type $\alpha \in [0, \underline{\alpha})$, we show that the payoff from action d exceeds the payoff from action nd at each of the three terminal nodes in Figure 2, working from top to bottom.

1. $M - P - \alpha D > 0$ by assumption
2. $M - P - \alpha D > 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner’s payoff from choosing action nd is $M - P$ and from d is $M - P - \alpha D + \rho(M - P)$. The homeowner will choose d if and only if

$$M - P - \alpha D + \rho(M - P) > M - P$$

$$\text{that is } \Leftrightarrow \alpha < \frac{\rho(M - P)}{D},$$

which is true because in this case $\alpha \in [0, \underline{\alpha})$ and $\underline{\alpha} = \frac{\rho(M - P)}{D}$.

Knowing that homeowners with $\alpha \in [0, \underline{\alpha})$ always choose action d , the lender will choose action na because her payoff from doing so, $P - F$, strictly exceeds her payoff from offering a , $P - F - \rho(M - P)$.

By backward induction, knowing that the lender will choose na , the homeowner will be indifferent between choosing s and ns at the initial node because the payoff is $M - P - \alpha D$ in each case.

For homeowners of type $\alpha \in [\underline{\alpha}, \bar{\alpha})$, we show that the payoff from action d exceeds the payoff from action nd at the top two terminal nodes and the payoff from nd exceeds the payoff from d at the bottom terminal node:

1. $M - P - \alpha D > 0$ by assumption
2. $M - P - \alpha D > 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P$ and from d is $M - P - \alpha D + \rho(M - P)$. The homeowner will choose d if and only if

$$\begin{aligned} M - P - \alpha D + \rho(M - P) &> M - P \\ \Leftrightarrow \alpha &< \frac{\rho(M - P)}{D}, \end{aligned}$$

which is false because in this case $\alpha \in [\underline{\alpha}, \bar{\alpha})$ and $\underline{\alpha} = \frac{\rho(M - P)}{D}$.

Knowing that homeowners with $\alpha \in [\underline{\alpha}, \bar{\alpha})$ choose action $nd|A = M - P$ and d otherwise, the lender will choose action a because her payoff from doing so, P , strictly exceeds her payoff from na , $P - F$. By backward induction, knowing that the lender will choose a , the homeowner will choose s at the initial node because $M - P > M - P - \alpha D$.

For homeowners of type $\alpha \in [\bar{\alpha}, 1]$, we show that the payoff from action nd exceeds the payoff from action d at each terminal node in Figure 2, working from top to bottom.

1. $M - P - \alpha D < 0$ by assumption
2. $M - P - \alpha D < 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P$ and from d is $M - P - \alpha D + \rho(M - P)$. The homeowner will choose d if and only if

$$\begin{aligned} M - P - \alpha D + \rho(M - P) &> M - P \\ \text{that is } \Leftrightarrow \alpha &< \frac{\rho(M - P)}{D}, \end{aligned}$$

which is false because in this case $\alpha \in [\bar{\alpha}, 1]$ and $\bar{\alpha} = \frac{\rho(M - P)}{D} > \frac{\rho(M - P)}{D}$.

Knowing that homeowners with $\alpha \in [\bar{\alpha}, 1]$ always choose action nd , the lender will choose action na because her payoff from doing so, M , strictly exceeds her payoff from offering a , P . By backward induction, knowing that the lender will choose na , the homeowner will be indifferent between choosing s and ns at the initial node because the payoff is 0 in each case. ■

Proof of Proposition 2

Proof. We show that the strategy profile is a subgame perfect Nash equilibrium by solving the game in Figure 3 by backward induction. The assumption that $\rho(M - P) \geq (1 - \rho)I_2$ ensures that $\underline{\alpha}' \in [0, \underline{\alpha}]$.

For homeowners of type $\alpha \in [0, \underline{\alpha}')$, we show that the payoff from action d exceeds the payoff from action nd at each terminal node in Figure 3, working from top to bottom.

1. $M - P - \alpha D > 0$ by assumption
2. $M - P - \alpha D > 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P + I_2$ and from d is $M - P - \alpha D + \rho(M - P + I_2)$. The homeowner will choose d if and only if

$$M - P - \alpha D + \rho(M - P + I_2) > M - P + I_2$$

$$\text{that is } \Leftrightarrow \alpha < \frac{\rho(M - P) - (1 - \rho)I_2}{D},$$

which is true because in this case $\alpha \in [0, \underline{\alpha}')$ and $\underline{\alpha}' = \frac{\rho(M - P) - (1 - \rho)I_2}{D}$.

Knowing that homeowners with $\alpha \in [0, \underline{\alpha}')$ always choose action d , the lender will compare her payoff from a , which is $P - F - \rho(M - P - I_2) + I_1$, to her payoff from choosing action na , which is $P - F$. The lender will choose a if and only if

$$P - F - \rho(M - P - I_2) + I_1 \geq P - F$$

$$\text{that is, } \Leftrightarrow I_1 \geq \rho(M - P - I_2),$$

which is false by assumption. Hence the lender will choose na . By backward induction, knowing that the lender will choose na , the homeowner will be indifferent between choosing s and ns at the initial node because the payoff is $M - P - \alpha D$ in either case.

For homeowners of type $\alpha \in [\underline{\alpha}', \bar{\alpha})$, we show that the payoff from action d exceeds the payoff from action nd at the top two terminal

nodes and the payoff from nd exceeds the payoff from d at the bottom terminal node:

1. $M - P - \alpha D > 0$ by assumption
2. $M - P - \alpha D > 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P + I_2$ and from d is $M - P - \alpha D + \rho(M - P + I_2)$. The homeowner will choose d if and only if

$$\begin{aligned} M - P - \alpha D + \rho(M - P + I_2) &> M - P + I_2 \\ \Leftrightarrow \alpha &< \frac{\rho(M - P) - (1 - \rho)I_2}{D}, \end{aligned}$$

which is false because in this case $\alpha \in [\underline{\alpha}', \bar{\alpha})$ and $\underline{\alpha}' = \frac{\rho(M - P) - (1 - \rho)I_2}{D}$.

Knowing that homeowners with $\alpha \in [\underline{\alpha}', \bar{\alpha})$ choose action $nd|A = M - P$ and d otherwise, the lender will choose action a because her payoff from doing so, $P + I_1 + I_2$, strictly exceeds her payoff from na , $P - F$. By backward induction, knowing that the lender will choose a , the homeowner will choose s at the initial node because $M - P + I_2 > M - P - \alpha D$.

For homeowners of type $\alpha \in [\bar{\alpha}, 1]$, we show that the payoff from action nd exceeds the payoff from action d at each terminal node in Figure 2, working from top to bottom.

1. $M - P - \alpha D < 0$ by assumption
2. $M - P - \alpha D < 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P + I_2$ and from d is $M - P - \alpha D + \rho(M - P + I_2)$. The homeowner will choose d if and only if

$$\begin{aligned} M - P - \alpha D + \rho(M - P + I_2) &> M - P + I_2 \\ \text{that is } \Leftrightarrow \alpha &< \frac{\rho(M - P) - (1 - \rho)I_2}{D}, \end{aligned}$$

which is false because in this case $\alpha \in [\bar{\alpha}, 1]$ and $\bar{\alpha} = \frac{\rho(M - P) - (1 - \rho)I_2}{D} > \frac{\rho(M - P) - (1 - \rho)I_2}{D}$.

Knowing that homeowners with $\alpha \in [\bar{\alpha}, 1]$ always choose action nd , the lender will compare her payoff from a , which is $P + I_1 + I_2$, to her

payoff from choosing action na which is M . The lender will choose a if and only if

$$P + I_1 + I_2 \geq M,$$

$$\text{that is } \Leftrightarrow I_1 + I_2 \geq M - P,$$

which is false by assumption. Thus the lender will choose na in this case. By backward induction, knowing that the lender will choose na , the homeowner will be indifferent between choosing s and ns at the initial node because his payoff is 0 in either case. ■

Proof of Proposition 3

Proof. We show that the strategy profile is a subgame perfect Nash equilibrium by solving the game in Figure 3 by backward induction. The assumption that $\rho(M - P) \geq (1 - \rho)I_2$ ensures that $\underline{\alpha}' \in [0, \underline{\alpha})$.

For homeowners of type $\alpha \in [0, \underline{\alpha}')$, we show that the payoff from action d exceeds the payoff from action nd at each terminal node in Figure 3, working from top to bottom.

1. $M - P - \alpha D > 0$ by assumption
2. $M - P - \alpha D > 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P + I_2$ and from d is $M - P - \alpha D + \rho(M - P + I_2)$. The homeowner will choose d if and only if

$$M - P - \alpha D + \rho(M - P + I_2) > M - P + I_2$$

$$\text{that is } \Leftrightarrow \alpha < \frac{\rho(M - P) - (1 - \rho)I_2}{D},$$

which is true because in this case $\alpha \in [0, \underline{\alpha}')$ and $\underline{\alpha}' = \frac{\rho(M - P) - (1 - \rho)I_2}{D}$.

Knowing that homeowners with $\alpha \in [0, \underline{\alpha}')$ always choose action d , the lender will compare her payoff from a , which is $P - F - \rho(M - P - I_2) + I_1$, to her payoff from choosing action na , which is $P - F$. The lender will choose a if and only if

$$P - F - \rho(M - P - I_2) + I_1 \geq P - F$$

$$\text{that is, } \Leftrightarrow I_1 \geq \rho(M - P - I_2),$$

which is true by assumption. Hence the lender will choose a . By backward induction, knowing that the lender will choose a , the homeowner

will compare choosing ns with choosing s . He will choose the latter if and only if

$$M - P - \alpha D + \rho(M - P + I_2) \geq M - P - \alpha D,$$

which is true. Hence the homeowner will indeed choose s .

For homeowners of type $\alpha \in [\underline{\alpha}', \bar{\alpha})$, we show that the payoff from action d exceeds the payoff from action nd at the top two terminal nodes and the payoff from nd exceeds the payoff from d at the bottom terminal node:

1. $M - P - \alpha D > 0$ by assumption
2. $M - P - \alpha D > 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P + I_2$ and from d is $M - P - \alpha D + \rho(M - P + I_2)$. The homeowner will choose d if and only if

$$\begin{aligned} M - P - \alpha D + \rho(M - P + I_2) &> M - P + I_2 \\ \Leftrightarrow \alpha &< \frac{\rho(M - P) - (1 - \rho)I_2}{D}, \end{aligned}$$

which is false because in this case $\alpha \in [\underline{\alpha}', \bar{\alpha})$ and $\underline{\alpha}' = \frac{\rho(M - P) - (1 - \rho)I_2}{D}$.

Knowing that homeowners with $\alpha \in [\underline{\alpha}', \bar{\alpha})$ choose action $nd|A = M - P$ and d otherwise, the lender will choose action a because her payoff from doing so, $P + I_1 + I_2$, strictly exceeds her payoff from na , $P - F$. By backward induction, knowing that the lender will choose a , the homeowner will choose s at the initial node because $M - P + I_2 > M - P - \alpha D$.

For homeowners of type $\alpha \in [\bar{\alpha}, 1]$, we show that the payoff from action nd exceeds the payoff from action d at each terminal node in Figure 3, working from top to bottom.

1. $M - P - \alpha D < 0$ by assumption
2. $M - P - \alpha D < 0$ by assumption
3. Given that the lender is offering $A = M - P$, the homeowner's payoff from choosing action nd is $M - P + I_2$ and from d is $M - P - \alpha D + \rho(M - P + I_2)$. The homeowner will choose d if and only if

$$\begin{aligned} M - P - \alpha D + \rho(M - P + I_2) &> M - P + I_2 \\ \text{that is } \Leftrightarrow \alpha &< \frac{\rho(M - P) - (1 - \rho)I_2}{D}, \end{aligned}$$

which is false because in this case $\alpha \in [\bar{\alpha}, 1]$ and $\bar{\alpha} = \frac{(M-P)}{D} > \frac{\rho(M-P)-(1-\rho)I_2}{D}$.

Knowing that homeowners with $\alpha \in [0, \underline{\alpha})$ always choose action nd , the lender will compare her payoff from a , which is $P + I_1 + I_2$, to her payoff from choosing action na , which is M . The lender will choose a if and only if

$$P + I_1 + I_2 \geq M$$

that is $\Leftrightarrow I_1 + I_2 \geq M - P$,

which is false by assumption. Thus the lender will choose na in this case. By backward induction, knowing that the lender will choose na , the homeowner will be indifferent between choosing s and ns at the initial node because the payoff from either action is 0. ■

REFERENCES

- Adelino, Manuel, Kristopher Gerardi, and Paul S. Willen. 2013. "Why Don't Lenders Renegotiate More Home Mortgages? Redefaults, Self-cures, and Securitization." *Journal of Monetary Economics* 60 (October): 835–53.
- Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Thomasz Piskorski, and Amit Seru. 2012. "Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program." Working Paper 18311. Cambridge, Mass.: National Bureau of Economic Research. (August).
- Ambrose, Brent W., and Charles A. Capone Jr.. 1996. "Cost-Benefit Analysis of Single-Family Foreclosure Alternatives." *Journal of Real Estate Finance and Economics* 13 (September): 105–20.
- Andersson, Fredrik, Souphala Chomsisengphet, Dennis Glennon, and Feng Li. 2013. "The Changing Pecking Order of Consumer Defaults." *Journal of Money, Credit and Banking* 45 (March/April): 251–75.
- Campbell, John Y., and João F. Cocco. 2015. "A Model of Mortgage Default." *Journal of Finance* 70 (August): 1495–554.
- Campbell, John Y., Stefano Giglio, and Parag Pathak. 2011. "Forced Sales and House Prices." *American Economic Review* 101 (August): 2108–31.
- Danis, Michelle A., and Anthony Pennington-Cross. 2008. "The Delinquency of Subprime Mortgages." *Journal of Economics and Business* 60 (January/February): 67–90.
- Deng, Yongheng, John M. Quigley, and Robert Van Order. 2000. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica* 68 (March): 275–307.
- Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. 2010. "What 'Triggers' Mortgage Default?" *American Economic Review* 100 (May): 490–4.
- Foote, Christopher L., Kristopher Gerardi, Lorenz Goette, and Paul S. Willen. 2010. "Reducing Foreclosures: No Easy Answers." In *NBER Macroeconomics Annual*, vol. 24, ed. Daron Acemoglu, Kenneth Rogoff, and Michael Woodford. Chicago: University of Chicago Press, 89–138.

- Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen. 2008. "Negative Equity and Foreclosure: Theory and Evidence." *Journal of Urban Economics* 64 (September): 234–45.
- Gerardi, Kristopher, Kyle F. Herkenhoff, Lee E. Ohanian, and Paul S. Willen. 2013. "Unemployment, Negative Equity and Strategic Default." Federal Reserve Bank of Atlanta Working Paper 2013-4 (August).
- Gerardi, Kristopher and Wenli Li. 2010. "Mortgage Foreclosure Prevention Efforts." Federal Reserve Bank of Atlanta *Economic Review* 95 (October): 1–13.
- Gerardi, Kristopher, Adam Hale Shapiro, and Paul S. Willen. 2008. "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures." Federal Reserve Bank of Boston Working Paper 07-15 (May).
- Gerardi, Kristopher, and Paul S. Willen. 2009. "Subprime Mortgages, Foreclosures, and Urban Neighborhoods." *BE Journal of Economic Analysis & Policy* 9 (March): 1–37.
- Ghent, Andra C. June 2011. "Securitization and Mortgage Renegotiation: Evidence from the Great Depression." *Review of Financial Studies* 24 (June): 1814–47.
- Ghent, Andra C., and Marianna Kudlyak. 2011. "Recourse and Residential Mortgage Default: Evidence from US States." *Review of Financial Studies* 24 (September): 3139–86.
- Goodman, Laurie S., Roger Ashworth, Brian Landy, and Ke Yin. 2010. "Negative Equity Trumps Unemployment in Predicting Defaults." *Journal of Fixed Income* 19 (Spring): 67–72.
- Hembre, Erik. 2014. "HAMP, Home Attachment, and Mortgage Default." Working Paper (February).
- Making Home Affordable Program. 2010. "Making Home Affordable Program Handbook for Servicers of Non-GSE Mortgages Version 2.0." https://www.hmpadmin.com/portal/programs/docs/hamp_servicer/mhahandbook_20.pdf [August 10, 2015].
- Mayer, Christopher, Edward Morrison, and Tomasz Piskorski. 2009. "A New Proposal for Loan Modifications." *Yale Journal on Regulation* 26 (Summer): 417–29.
- Mayer, Christopher, Edward Morrison, Tomasz Piskorski, and Arpit Gupta. 2014. "Mortgage Modification and Strategic Behavior: Evidence from a Legal Settlement with Countrywide." *American Economic Review* 104 (September): 2830–57.

- Neelakantan, Urvi, Kimberly Zeuli, Shannon McKay, and Nika Lazaryan. 2012. "Staring Down Foreclosure: Findings from a Sample of Homeowners Seeking Assistance." Manuscript, Federal Reserve Bank of Richmond.
- RealtyTrac Staff. 2014. "1.4 Million U.S. Properties with Foreclosure Filings in 2013 Down 26 Percent to Lowest Annual Total Since 2007." RealtyTrac.com (January 13).
- Scharlemann, Therese C., and Stephen H. Shore. 2016. "The Effect of Negative Equity on Mortgage Default: Evidence from HAMP's Principal Reduction Alternative." *Review of Financial Studies* 29 (October): 2850–83.
- U.S. Department of the Treasury. 2014. "HAMP Application Activity by Servicer, as of February 28, 2014." <http://www.treasury.gov/initiatives/financial-stability/reports/Pages/HAMP-Servicer.aspx> [August 10, 2015].
- Wang, Ko, Leslie Young, and Yuqing Zhou. 2002. "Nondiscriminating Foreclosure and Voluntary Liquidating Costs." *Review of Financial Studies* 15 (April): 959–85.
- White, Alan M. 2009a. "Deleveraging the American Homeowner: The Failure of 2008 Voluntary Mortgage Contract Modifications." *Connecticut Law Review* 41 (May): 1107–31.
- White, Alan M. 2009b. "Rewriting Contracts, Wholesale: Data on Voluntary Mortgage Modifications from 2007 and 2008 Remittance Reports." *Fordham Urban Law Journal* 36: 509–35.