

# The Time-Varying Price of Financial Intermediation in the Mortgage Market

Andreas Fuster, Stephanie H. Lo, and Paul S. Willen

## Abstract

The U.S. mortgage market links homeowners with savers all over the world. In this paper, we ask how much of the flow of money from savers to borrowers actually goes to the intermediaries that facilitate these transactions. Based on a new methodology and a new administrative dataset, we find that the price of intermediation, measured as a fraction of the loan amount at origination, is large—142 basis points on average over the 2008–2014 period. At daily frequencies, intermediaries pass on the price changes in the secondary market to borrowers in the primary almost completely. At monthly frequencies, the price of intermediation fluctuates significantly and is highly sensitive to volume, likely reflecting capacity constraints: a one standard deviation increase in applications for new mortgages leads to a 30–35 basis point increase in the price of intermediation. Additionally, over 2008–2014, the price of intermediation increased about 30 basis points per year, potentially reflecting higher mortgage servicing costs and an increased legal and regulatory burden. Taken together, the sensitivity to volume and the positive trend led to an implicit total cost to borrowers of about \$135 billion over this period. Finally, the increases in application volume associated with “quantitative easing” (QE) led to substantial increases in the price of intermediation, which attenuated the benefits of QE to borrowers.

**JEL Classifications:** E44, E52, G21, L11

**Keywords:** mortgage finance, financial intermediation, monetary policy transmission

---

Andreas Fuster is a research officer in the Capital Markets Function at the Federal Reserve Bank of New York. His e-mail address is [andreas.fuster@ny.frb.org](mailto:andreas.fuster@ny.frb.org). At the time this paper was written, Stephanie H. Lo was a Ph.D. candidate in the department of economics at Harvard University and a graduate fellow in the research department at the Federal Reserve Bank of Boston. Her e-mail address is [shlo@fas.harvard.edu](mailto:shlo@fas.harvard.edu). Paul S. Willen is a senior economist and policy advisor in the research department at the Federal Reserve Bank of Boston and a research associate at the National Bureau of Economic Research. His e-mail address is [paul.willen@bos.frb.org](mailto:paul.willen@bos.frb.org).

The authors thank John Campbell, Andy Davidson, Fernando Ferreira, Paul Goldsmith-Pinkham, Sam Hanson, Sean Hundtofte, Dwight Jaffee, Michael Reher, Adi Sunderam, Stijn Van Nieuwerburgh, and Alex Zevelev, plus seminar audiences at Brandeis, the Homer Hoyt Institute, the Federal Reserve Banks of New York and St. Louis, the Board of Governors, the NBER Summer Institute, the NYC Real Estate conference at Baruch, and the Universities of St. Gallen and Zurich for helpful comments and suggestions.

This paper presents preliminary analysis and results intended to stimulate discussion and critical comment. The views expressed herein are those of the authors and do not indicate concurrence by the Federal Reserve Bank of Boston, or by the principals of the Board of Governors, or the Federal Reserve System. This paper, which may be revised, is available on the web site of the Federal Reserve Bank of Boston at <http://www.bostonfed.org/economic/wp/index.htm>.

**This version: August, 2017; original version posted January, 2017**

# 1 Introduction

Mortgage lending is one of the main activities of the U.S. financial sector and a principal driver of its growth in recent decades (Greenwood and Scharfstein, 2013). For most U.S. mortgage transactions, intermediaries directly connect borrowers with capital market investors through the market for mortgage-backed securities (MBS), rather than funding loans from deposits or other sources.<sup>1</sup> One might expect that the price of this service should be fairly low and stable. Instead, we show that over the 2008–2014 period, the price of intermediation was high and volatile. We explore the drivers of this variation and study its implications for the pass-through of monetary policy, specifically the Federal Reserve’s large-scale asset purchase programs, colloquially known as “quantitative easing” (QE).

Our first contribution is to develop a new methodology to measure the *price of intermediation* in the mortgage market, which we define as the payment made to the intermediary for linking the borrower with the ultimate supplier of funds for the loan. This payment covers costs associated with originating, underwriting, and servicing the loan and also may include a component representing profits. A way to think about the role of the intermediary is as a dealer who buys a loan from the borrower (meaning the right to a stream of principal and interest payments), and sells it to investors in the form of an MBS. The price of intermediation is then the the intermediary’s dollar margin. The payment to the borrower includes not only the loan’s principal balance but also a “rebate,” an upfront payment made by the intermediary to the borrower.<sup>2</sup> The rebate depends on the loan’s interest rate and for a given interest rate typically fluctuates on a daily basis.

For our analysis, the two central inputs thus are the value of a mortgage in the secondary market and the rebate the intermediary pays to the borrower. Measuring the price of a loan in the MBS market is relatively straightforward. It is much more challenging to measure the rebates, since in general, the “rate sheets” on which the rebates are quoted are not publicly disclosed. A key innovation in our analysis is that we use a new administrative dataset that documents these rebates on a daily basis as a function of the interest rate and the characteristics of the borrower and the loan. The data come from Optimal Blue, an industry-leading provider of real-time data to loan officers.

We use our new methodology for measuring the price of intermediation for several complementary analyses. First, we study the high-frequency (daily) pass-through of MBS price

---

<sup>1</sup>Over our sample period (2008–2014), about 80 percent of new mortgage loans are securitized through the “agency” MBS market, where the government-backed agencies Fannie Mae, Freddie Mac, and Ginnie Mae insure the timely payment of principal and interest payments to investors.

<sup>2</sup>Rebates are typically positive but can also be negative (i.e. the payment is from the borrower to the intermediary) in which case they are known as “discount points.”

changes to rebates over the 2008–2014 period, initially focusing on six major monetary policy announcements (mostly QE-related) and then looking across all the days in our sample period. Next, we document how the price of intermediation evolved at a lower (monthly) frequency, explore the potential drivers of the variation, and study the implied costs to borrowers. Finally, we clarify the relationship between our price of intermediation and a commonly used measure that compares mortgage rates to yields in the MBS market.

Our main findings are as follows: The pass-through of MBS price changes to the consumer mortgage market following QE announcements was generally large, though it varied with the level of demand for mortgages (measured by daily applications for new loans) at the time of the announcement. When demand was strong, such as around the expansion of “QE1” in March 2009, intermediaries passed on a smaller share of the secondary market price increases to borrowers. More generally, we find that at the daily frequency, rebates move very closely with prices in the MBS market. We estimate that, on average, a one dollar change in the MBS price leads to a 92 cent change in the rebate paid to the borrower. Pass-through, while high, is not symmetric: a one dollar decrease in the MBS price translates into a one dollar reduction in the rebate, but a one dollar increase only leads to an 80 cent rebate increase. Importantly, high levels of demand substantially aggravate this asymmetry in pass-through.

Our findings of strong high-frequency pass-through are in stark contrast to other consumer credit markets. [Ausubel \(1991\)](#) showed that over 1983–1988, credit card interest rates did not respond at all to changes in banks’ cost of funds, and other work has confirmed this finding in more recent data (e.g., [Calem, Gordy, and Mester, 2006](#)). Evidence from interest rates on deposits paints a similar picture. For example, [Driscoll and Judson \(2013\)](#) show that depository institutions change interest rates on short-term certificates of deposit (CDs) roughly every six to seven weeks and that the coefficient on the fed funds rate in a regression of changes in CD rates is only 0.4 at weekly frequencies.<sup>3</sup> In contrast, we find that in the mortgage market lenders change rates at least once a day and a comparable regression generates a coefficient of 0.92. Thus, mortgage pricing appears much more tightly linked to capital markets than other consumer credit; this could be due to the particular market structure that features brokers and loan officers comparing offers across multiple intermediaries.

Despite the strong daily pass-through, the price of intermediation is highly variable over time. In particular, it strongly increases with the volume of new mortgage applications: at a monthly frequency, a one standard deviation increase in application volume is associated with a 30–35 basis point increase in the price of intermediation (relative to an average of 142

---

<sup>3</sup>See also [Hannan and Berger \(1991\)](#), [Neumark and Sharpe \(1992\)](#), [Yankov \(2014\)](#) and [Drechsler, Savov, and Schnabl \(2015\)](#) for related evidence on CDs/deposit rates. [Mora \(2014\)](#) documents similarly low pass-through in auto loan rates.

basis points over our sample period). We show that demand increases cause higher prices in this market using an instrumental variable strategy, exploiting the fact that changes in the gap between rates on outstanding mortgages and market interest rates should affect the demand for intermediation (by changing borrowers' incentive to refinance their loans) but not the supply, which instead should reflect the costs of underwriting and servicing mortgages. Our findings are consistent with the existence of significant capacity constraints in the mortgage industry, for which we also present direct evidence based on the processing time of new loan applications.

We also document a significant upward trend in the price of intermediation between 2009 and 2014; using our baseline specification, this price increase amounts to 30 basis points per year. This trend is not explained by a change in interest rate volatility (a driver of intermediaries' hedging costs) or a change in market concentration. Per-worker labor costs in the real estate credit sector are correlated with the price of intermediation and account for a significant share of the increase in regressions but it is difficult to know which way the causality runs. Even after controlling for labor costs, however, a positive trend remains, bearing out anecdotal evidence of an increased legal and regulatory burden leading to increased intermediation costs over this period. Consistent with this interpretation, we show that an important part of the upward price trend appears to be due to a decrease in intermediaries' valuation of mortgage servicing rights (relative to the value of the cash flows in the MBS market), which may reflect increased loan servicing costs and the changed treatment of servicing rights under revised capital regulations.

The variation and especially the upward trend in the price of intermediation have resulted in substantial cost increases for U.S. mortgage borrowers. Specifically, over 2009–2014, had the price of intermediation been insensitive to loan application volume and had there been no upward trend in this price, borrowers collectively would have received roughly an additional \$135 billion (in present value terms), holding the volume and timing of originations constant. Our methodological approach allows us to translate the increase in the price of intermediation into its effect on interest rates for borrowers, holding the upfront cost to the borrower constant. We find that over extended intervals during our sample period, mortgage rates would have been 30–40 basis points lower under the counterfactual described above.

Our work relates to various strands of existing research. Most broadly, we contribute to the voluminous literature studying how monetary policy and financial market conditions affect the real economy through the credit channel (e.g. [Bernanke, Gertler, and Gilchrist, 1999](#); [Bernanke and Gertler, 1995](#); [Kashyap and Stein, 2000](#)). Specifically, we focus on the link between financial markets (which in turn may be affected by monetary policy) and the borrowing rates faced by households. Recent contributions with a similar focus include

Gertler and Karadi (2015) and Gilchrist, López-Salido, and Zakrajšek (2015). While other work has emphasized time-varying risk-premia—for instance in corporate bonds (Gilchrist and Zakrajšek, 2012)—we focus specifically on changes in mortgage interest rates due to the time-varying margins of intermediaries. Prior research suggests that mortgage interest rates in turn affect the real economy; for instance, Walentin (2014) finds that mortgage rate spreads have significant explanatory power for several macro variables. Malamud and Schrimpf (2017) study theoretically how intermediation markups interact with monetary policy transmission.

A growing literature has studied the effects of QE on lending rates and volumes. Most directly, this paper builds on Fuster and Willen (2010) who looked at QE1 announcements in more detail based on an earlier version of the rate sheet data we use here. Hancock and Passmore (2011, 2015) also study the impact of unconventional monetary policy on consumer mortgage rates, while Stroebel and Taylor (2012), Krishnamurthy and Vissing-Jorgensen (2011, 2013) and Boyarchenko, Fuster, and Lucca (2015) focus primarily on MBS market spreads. The effects of QE announcements on the quantity of mortgage originations (especially refinancings) are studied by Fuster and Willen (2010), Beraja et al. (2015), and Di Maggio, Kermani, and Palmer (2016), who respectively focus on heterogeneity across borrower types, regions, and market segments.

In the banking literature, Hannan and Berger (1991), Neumark and Sharpe (1992) and Drechsler, Savov, and Schnabl (2015) have linked the slow adjustment of deposit rates to market power. In the mortgage market, Scharfstein and Sunderam (2013) study heterogeneity across counties in the sensitivity of rates and refinancing to changes in MBS yields, finding less pass-through in more concentrated markets. We study whether changes in market concentration could explain the increase in intermediation prices that we document, but since concentration in mortgage lending fell substantially over the 2010–2014 period, this does not appear to be a promising avenue for explaining the patterns in our data.

Related to our focus on capacity constraints in the mortgage market, Sharpe and Sherlund (2016) present evidence consistent with limited capacity affecting the types of loans that lenders choose to originate. Fuster et al. (2013), who focus on longer-term trends in intermediary margins (and who did not use the high-quality daily rate sheet data used here) also highlight capacity constraints as a potentially important explanation of temporarily high margins. Other work has focused on the capacity constraints of mortgage servicers (who may be the same entity as the originators we study) and how these constraints have affected modification activity during the crisis (e.g. Cordell et al., 2009; Maturana, 2015).

Our argument that the upward trend in the price of intermediation partly reflects an increased legal and regulatory burden is also consistent with other recent changes in the

structure of the mortgage market. In particular, [Buchak et al. \(2017\)](#) study the drivers behind the rapid increase in market share of non-depository institutions (shadow banks) in mortgage originations, and find that this is to an important extent due to regulatory factors.

Finally, studying the market for financial intermediation in mortgages is interesting from an industrial organization perspective. Our analysis parallels other work that studies how product prices react to changes in input prices—for instance, how the price of gasoline reacts to the price of crude oil (e.g. [Borenstein, Cameron, and Gilbert, 1997](#); [Bachmeier and Griffin, 2003](#)). We do find some asymmetry in the pass-through, with positive MBS price changes being more slowly reflected in consumer market prices than negative changes, in line with existing findings of similar asymmetries in many markets (e.g. [Peltzman, 2000](#)).

## 2 Financial Intermediation in the Mortgage Market

In this section we provide a brief overview of intermediation in the mortgage market through the lens of a simple model; see e.g. [Fuster et al. \(2013\)](#) for a more extensive discussion of the institutional details.

### 2.1 A Simple Model of the Mortgage Market

Consider a simple model of the market for financial intermediation in mortgages, where we define financial intermediation as the service of matching a borrower with an MBS investor. We follow industry parlance and refer to the consumer and MBS markets respectively as the primary and secondary markets. We start with a lender/intermediary who makes a loan to a borrower with a fixed interest rate  $r^n$ . A central feature of the U.S. mortgage market is that, in addition to providing the borrower with the principal on the loan, the intermediary also pays a “rebate” to the borrower to cover closing costs and other expenses.<sup>4</sup> This upfront payment, which goes by many different names (yield spread premium or YSP, service release premium, (negative) discount points), plays a central role in all mortgage transactions but is often not explicitly disclosed to the borrower who just sees the rebate in the form of changed closing costs.

The size of that rebate, denoted  $YSP(r^n)$ , depends on the interest rate on the loan, known in the industry as the note rate. A higher note rate is more valuable to the lender since it generates higher future cash flows and results in a higher rebate. One way to think about this transaction is that in originating a mortgage, the intermediary *buys* the mortgage

---

<sup>4</sup>The rebate can be negative (i.e. the borrower pays the intermediary) in which case it is referred to as the borrower “paying (discount) points.”

from the borrower by paying the principal plus the rebate. That is, the price paid for a mortgage with \$100 principal and note rate  $r^n$  is:

$$p_{YSP}^n = 100 + YSP(r^n).$$

After the intermediary buys the loan from the borrower, it turns around and sells it to investors. In general, intermediaries can sell loans to many different types of investors but over the period we study, about 80 percent of new mortgage lending was funded through “agency” MBS guaranteed by the government-sponsored enterprises (GSEs) Fannie Mae and Freddie Mac or the government agency Ginnie Mae; only about 20 percent of loans were kept on bank balance sheets.<sup>5</sup> Our discussion will focus on a loan sold through a MBS issued by Fannie Mae; the mechanics for Freddie Mac or Ginnie Mae are very similar.

To sell loans in the secondary market, the intermediary typically starts by putting together a pool of loans and then exchanges the pool with Fannie Mae for a MBS collateralized by those loans. The intermediary pays a monthly premium called a guarantee fee (g-fee) and a one-time upfront fee called the “loan-level price adjustment” (LLPA) to Fannie Mae, in exchange for which Fannie Mae ensures timely payment of principal and interest. After the swap with Fannie Mae, the intermediary holds a tradeable MBS with cash flows based on the payments from the underlying mortgages. The intermediary can then sell the MBS to an investor (such as a pension fund or a central bank) in the secondary market.

The process of closing, packaging the loan into a MBS, and delivering it to an investor takes several weeks but, fortunately for the intermediary, there is a highly liquid forward market, the “to-be-announced” or TBA market, for agency MBS. More or less immediately after agreeing to terms—“locking” the loan—with the borrower, the intermediary can sell the loan forward in the TBA market. We define  $p_{TBA}^n$  to be the price of a loan in the TBA market (per \$100 principal) net of the anticipated payments to Fannie Mae associated with selling, securitizing and insuring the loan. Since we study the market for intermediation, and not the market for mortgages *per se*, we take  $p_{TBA}^n$  as exogenously given.

We then define  $\phi^n$ , the market price of intermediation for a loan with note rate  $r^n$  as:

$$\phi^n \equiv p_{TBA}^n - p_{YSP}^n. \tag{1}$$

Our empirical analysis will investigate how  $\phi^n$ , the price of intermediation, changes over time and in response to economic shocks. To gain an intuitive sense of how this happens, consider a profit-maximizing and (for simplicity) price-taking intermediary that originates

---

<sup>5</sup>See e.g. p.8 of <http://www.urban.org/UploadedPDF/2000150-Housing-Finance-At-A-Glance.pdf>.

mortgages. We assume that they make  $q^n$  loans of note rate  $r^n$  with the objective:

$$\max_{q^n} \phi^n \cdot q^n - C(q^n),$$

where  $C(\cdot)$  includes the intermediary's costs of underwriting, originating, and servicing loans (as discussed in the next subsection) and delivering the resulting MBS to an investor. Trivially, the first-order condition is

$$C'(q^n) = \phi^n. \tag{2}$$

Equation (2) implies the usual condition that the price equals the marginal cost which, in a perfectly competitive market with free entry, would imply zero economic profit in equilibrium. We will discuss the market structure further below.

Figure 1 graphically shows the equilibrium in the market for intermediation. In the standard supply-and-demand representation of a market, the quantity  $q$  on the  $x$ -axis is the number of mortgages originated and the price  $\phi$  on the  $y$ -axis is the price of intermediation. If we assume increasing marginal costs, equation (2) implies that supply is upwardly sloping in quantity. Looking at the demand side, households demand fewer mortgages as the price of intermediation ( $\phi$ ) goes up. To see why, note that holding  $p_{TBA}^n$  constant, any increase in  $\phi$  reduces the rebate paid to the borrower for a given note rate  $r^n$ .

One goal of this paper is to understand how changes in the price of a mortgage in the secondary market translates into the prices a borrower faces in the primary market. Figure 1 shows that the effect of a change in  $p_{TBA}$  on  $\phi$  and  $q$  depends on the shape of the supply curve, which in turn depends on intermediaries' marginal costs. As depicted in the top panel, if the marginal cost of originating mortgages is increasing, then an increase in  $p_{TBA}$  will lead to an increase in  $\phi$ , and thus the pass-through to primary market prices ( $p_{YSP}$ ) will be incomplete. If instead, as shown in the bottom panel, the marginal cost is constant, the supply curve is flat and changes in TBA prices will completely pass through to rebates (meaning  $\phi$  remains constant). Also, note that in this case, the loan quantity response to a change in TBA prices is larger. Of course one can imagine intermediate cases, such as a kinked supply curve that is flat up to some "normal capacity" level and then starts to slope upwards. If so, the pass-through would be complete at relatively low levels of  $q$ , but diminish once  $q$  is to the right of the kink.

For simplicity, the above discussion has considered a single note rate. In reality, intermediaries typically offer many different note rates,  $n = 1, \dots, N$ , which differ in their secondary and primary market values. The first-order condition (2) will still hold for all  $n$ , and if the marginal costs of producing different note rates are equal,  $\phi^n$  should be the same for all  $n$ .

It is important to understand that the implicit price paid to the intermediary,  $\phi$ , is only



one portion of a borrower’s expenditure to get a mortgage. The borrower also faces other expenses including appraisal, title insurance, an application fee, and a commission paid to the loan officer. Those expenses (which may partly be covered by the rebate) are beyond the scope of our analysis as these costs are not paid to the intermediary but to third parties.

## 2.2 Structure of the Mortgage Market

In our discussion above, we described a generic intermediary that matches a borrower in the primary market with an investor in the secondary market. The reality is more complicated as there is enormous variation in institutional arrangements in the U.S. mortgage market. In the retail market, a borrower works with a loan officer (LO) employed by the intermediary and the intermediary underwrites and funds the loan. In the wholesale market, the LO is a broker who links borrowers with one of several intermediaries which, in turn, underwrite and fund the loan. Finally, a LO can work for a “correspondent” lender, meaning a lender that funds and originates loans like a retail lender but contracts to sell the loan to another intermediary more or less immediately after closing.

As a practical matter, most borrowers cannot tell if their LO is working as a broker or as a correspondent lender and many LOs do both, so the correspondent and wholesale segments are often lumped together as “third-party origination” or TPO. Our data comes from an on-line platform that allows LOs to choose from offers made by different intermediaries and, as a result, is only relevant for the TPO market. Historically, TPO has accounted for as much as 70 percent of origination volume; over the period of our study, the TPO share fell somewhat, from 56 percent in 2008 to 44 percent in 2014.<sup>6</sup>

Mortgage intermediation is highly fragmented. According to data compiled pursuant to the Home Mortgage Disclosure Act (HMDA), there were 7,062 active mortgage lenders in 2014 down from 7,923 in 2010.<sup>7</sup> The top four firms in the industry accounted for 31 percent of origination volume in 2014 but the market share of the fifth-largest firm was only 3 percent and the tenth-largest less than 2 percent, indicating a large number of small but active lenders. Given the size of the U.S. mortgage market, even small players are sizable companies—the tenth-largest lender originated 23 billion dollars worth of loans in 2014. Mortgage lenders typically operate in multiple states and a significant number operate nationally, meaning that local markets are also fragmented. For instance, according to HMDA data, more than 700 lenders were active in the Los Angeles MSA each year over the

---

<sup>6</sup>All market share data comes from [Inside Mortgage Finance \(2016\)](#). Whether the loan is retail or TPO has no bearing on whether the intermediary sells the loan in the secondary market. In the part of the market we study, most retail loans, just like most TPO loans, end up securitized in agency MBS.

<sup>7</sup>See <https://www.ffiec.gov/press/pr092211.htm> and <https://www.ffiec.gov/press/pr092215.htm> respectively.

2008–2014 period, with a comparable share for the top four firms to their national share.

In Section 2.1, we assumed perfect competition. Given the facts discussed above, how realistic is this assumption? On one hand, an industry consisting of a large number of small but non-trivial firms and relatively free entry into different markets does not match the assumptions of a Cournot-type framework. A more plausible imperfectly competitive model for the market is monopolistic competition, which still allows for economic profits to be made even with many firms (see e.g. [Gabaix et al., 2016](#)). In our context, one could think of intermediaries gaining market power with specific LOs due the LO having a preference for working with that firm. However, in the Optimal Blue data that we use, a typical LO receives offers on daily basis from between five and ten different intermediaries for more or less identical products, suggesting that intermediaries do compete for business from LOs.

Going forward, we focus on the perfectly competitive case with upward sloping costs as our leading explanation for the data. However, we cannot rule out alternative explanations, including intermediaries with market power over particular LOs. Also, it is important to stress that the assumption of perfect competition does not imply that we should not observe accounting profits in the mortgage market or that an increase in demand for intermediation will not increase returns for intermediaries. In particular, the marginal cost includes payments to all factors of production. If we think of intermediaries as owning capital, either physical like computer systems or non-physical like reputational capital, then an increase in demand for intermediation will lead to increases in intermediaries’ accounting profits.

### 2.3 The Intermediary Cost Function $C(q^n)$

What are the costs to the intermediary? Intermediaries play two functions in the life of the loan, underwriting/originating the loan, and then servicing the loan once it has been made.

The first component of origination costs are the direct costs associated with the process of underwriting the loan. An intermediary must employ loan officers to work with borrowers, underwriters to review applications, a compliance department to make sure that the loan officers and underwriters are fulfilling their legal obligations, and so on. Additional costs include rent, information technology expenses, and advertising.

The second component of origination costs involve various forms of risk management. In particular, mortgage originators actively engage in “pipeline hedging,” meaning that they hedge financial risks between the time a borrower “locks” a rate/rebate combination and the time the loan closes and is delivered to a MBS investor, which typically takes somewhere between 30 and 90 days. The main risk is that the fraction of loans that are actually

originated is lower than expected.<sup>8</sup> This tends to occur primarily when interest rates fall between the time of lock and origination, since many borrowers may either try to renegotiate or go to a different intermediary. Such an outcome is costly for originators if they forward-sold the loan in the MBS market, since an inability to meet that commitment will require them to buy back part of their committed volume at a higher price (since rates fell). To hedge this risk, originators typically either use “mortgage options” (options on TBA contracts) or “swaptions”; the cost of the hedge will increase with implied interest rate volatility.

Another risk management cost stems from the possibility that the intermediary will be forced to buy back a loan that the guaranteeing agency determines to be in violation of its underwriting guidelines. The volume of such “putbacks” became very large for the loan vintages that performed the worst during the financial crisis. However, for new loans, such risks can be mitigated through careful underwriting. Also related in part to putback risk, there is a cost of capital, since intermediaries that want to sell loans to the guaranteeing agencies are required to maintain a certain net worth (as a cushion against future liabilities).

In addition to underwriting the loan, the intermediary’s second main task is to service the loan, which consists of collecting monthly payments from the borrowers after the loan is made, and in case the borrower becomes delinquent, working out a loss mitigation strategy and/or foreclosing. Servicing generates costs but can also generate income. Servicers receive “float income” (coming from a delay between when payments are received from borrowers and when they are passed on to investors). In addition, the servicer also gains an opportunity to cross-sell financial products and has the inside track to refinance the loan.

Formally, we can decompose the cost of intermediation into one-time origination costs,  $C_o$ , and per-period servicing costs,  $C_s$ . The present value (PV) of the costs of intermediation is then:

$$C(q^n) = C_o(q^n) + PV(C_s(q^n)). \quad (3)$$

Aside from the timing, there is also a fundamental difference in the aggregate cost function for underwriting versus servicing. As we show later, the demand for underwriting fluctuates enormously. A drop in interest rates can lead to massive waves of new loan applications which can tax the limited resources of the industry. A similar problem does not occur with servicing since refinancing typically has no effect on the aggregate number of loans being serviced. As a result, we would expect  $C'_o$  to be upward sloping but we would not expect  $C'_s$  to be. Thus, if we assume that the per-period cost of servicing a loan is a constant  $c_s$ , we

---

<sup>8</sup>There is also the risk that a higher-than-expected fraction of loans end up being originated. In expectation, a positive fraction of loans “fall out” due to idiosyncratic events, such as the property appraisal not coming in sufficiently high or the borrower being unable to produce required documentation.

can then write the marginal cost of producing and servicing a loan as<sup>9</sup>

$$C'(q^n) = C'_o(q^n) + PV(c_s). \quad (4)$$

### 2.3.1 The Valuation of Servicing Rights

One important institutional detail in agency MBS is how the servicer gets compensated. For as long as a loan is open, a servicer receives a monthly payment equal to 25 basis points (annual) of the mortgage amount.<sup>10</sup> As discussed above, the servicer receives additional income (or potential benefits), but also incurs costs. Our baseline calculation of  $\phi$  does not separately measure these additional costs/benefits to the servicer, and simply values the 25 basis points of servicing based on the value of the cash flow in the MBS market. We can write (using the superscript “ $n - 0.25$ ” to mean a note rate of  $r^n - 0.25$ ):

$$\begin{aligned} \phi^n &\equiv p_{TBA}^n - p_{YSP}^n \\ &= p_{TBA}^{n-0.25} + (p_{TBA}^n - p_{TBA}^{n-0.25}) - p_{YSP}^n \\ &= p_{TBA}^{n-0.25} + 0.25 \cdot \underbrace{\frac{p_{TBA}^n - p_{TBA}^{n-0.25}}{0.25}}_{\equiv mult_{MBS}^n} - p_{YSP}^n, \end{aligned}$$

where  $mult_{MBS}^n$  is a valuation multiple that transforms an interest rate strip (here, generating 25 basis points annually) into a present value.

In our analysis, we also consider a measure of the price of intermediation that nets out the costs and indirect benefits of servicing. To do this, we make use of an alternative multiple to value servicing, provided by a firm called MIAC (Mortgage Industry Advisory Corporation) that specializes in estimating the market value of servicing rights. The MIAC multiples lead to an alternative measure of intermediary margins,  $\pi^n$ :

$$\pi^n = p_{TBA}^{n-0.25} + 0.25 \cdot mult_{MIAC}^n - p_{YSP}^n, \quad (5)$$

where  $\pi$  is different from  $\phi$  in that it incorporates the additional benefits from servicing, but also the per-period costs. If  $c_s$  is the per-period net cost of servicing, equations (2) and (4)

---

<sup>9</sup>Note that  $c_s$  can change across loans originated at different times, as our empirical analysis will suggest that it has. The assumption here is only that the marginal cost is independent of the quantity of originations at a given time.

<sup>10</sup>The entity servicing the loan is not always the same as the one that originated the loan. However, since the intermediary that originally owns the “mortgage servicing right” (MSR) gets compensated when transferring it (there is a relatively active market in which MSRs are traded), this does not affect our discussion.

imply that in equilibrium,  $\phi = C'_o(q^n) + PV(c_s)$  and  $\pi = C'_o(q^n)$ .<sup>11</sup> The difference between  $\phi$  and  $\pi$  is thus equal to  $PV(c_s)$ , the present value of the net cost of servicing.

### 3 Measurement

In this section, we describe our approach to calculating the components of the price of intermediation  $\phi$ , defined in equation (1). First, we discuss how to measure  $p_{Y_{SP}}^n$ , including how we choose  $r^n$ , the note rate of the mortgage. We then turn to the question of how to measure  $p_{TBA}^n$ , the price of a given mortgage in the secondary market.

#### 3.1 Measuring the Primary Market Price

Intermediaries publish daily rate sheets which are, essentially, a menu of combinations of note rates and rebates. Figure 2 shows an example. The upper left panel shows that if the borrower commits to close a 30-year fixed-rate mortgage within 30 days with a note rate of 4.625 percent, the lender will pay the borrower a 70 cent rebate per 100 dollars of principal. As mentioned already, rebates can be negative. If instead the borrower chooses a 4.25 percent rate, the rebate would be -130 cents per 100 dollars of principal — i.e. the borrower would pay the lender.

Our rate sheet data come from a company called Optimal Blue (formerly known as Loan-Sifter) which provides data on rates and rebates to loan officers (LOs) in the third-party origination segment of the market. Based on digitized rate sheets, Optimal Blue provides LOs with a search engine that allows the LO to enter the characteristics of the loan such as the loan-to-value (LTV) ratio, the borrower’s FICO score, and the loan amount, and see different intermediaries’ offers for such a loan. Importantly, these offers are binding on the intermediary. If the LO selects an offer and delivers a loan that satisfies the undewriting guidelines, the intermediary has an obligation to fund it at the chosen rate/rebate combination. Intermediaries typically issue the first rate sheet of the day around 10 a.m., and then revise it frequently over the course of the day.<sup>12</sup> Our data only includes a single snapshot taken at the end of the day, meaning we have the last rate sheet issued on any given day.

Each LO has access to a subset of intermediaries with which the LO has an agreement,

---

<sup>11</sup>The measure  $\pi$  is close to but distinct from the measure of “originator profits and unmeasured costs” (or OPUC) of Fuster et al. (2013), with the difference that  $\pi$  does not explicitly use the “best execution” MBS coupon, whereas OPUC does and  $\pi$  does not include origination points. On the other hand,  $\pi$  is based on more accurate and higher frequency data than OPUC.

<sup>12</sup>Intermediaries, also referred to as “investors” in industry parlance, generally adjust their rate sheets on at least a daily level. Of the 35,400 rate sheets captured in our sample, only 2300 (about 6.5 percent) show no change from the previous day.

and offers for an identical loan may differ across LOs for a given intermediary. We obtain offers either for a generic LO that was set up specifically for us (for data from October 2008 until September 2009) or for five different representative loan officers (since September 2009). Further information about the Optimal Blue dataset can be found in Table 1. The number of lenders over which Optimal Blue searches fluctuates in our sample. Part of this difference reflects the entry and exit of lenders from the TPO market; for example, Bank of America exited in 2011.<sup>13</sup> In addition, because our searches since 2009 are based on the profiles of actual loan officers, any change in the set of investors with whom a specific loan officer has a relationship may lead to a change in the number of offers in our sample. We do not know the identities of the individual lenders, but do know that essentially all the largest lenders are in the Optimal Blue data. Appendix A provides additional detail about the data, including how we aggregate across lenders.

Our baseline scenario involves a fixed-rate mortgage on an owner-occupied property located in Los Angeles, CA, a borrower with a FICO of 750, a term of 30 years, a loan amount of \$300,000, a loan-to-value ratio of 80 percent, no prepayment penalty, and a 30-day lock period. As discussed later, we consider alternative scenarios as well.

### 3.1.1 Tracking a Mortgage Note Rate Over Time

In principle, we could choose a single note rate and follow that note rate over our entire sample period. Unfortunately, in reality, the set of note rates quoted at a point in time is quite narrow and changes often as interest rates evolve, since there is no market for loans with either very high or very low rebates (high rebates are unattractive to lenders due to the prepayment option; low rebates are unattractive to borrowers because they require high upfront payments). As shown in Table 1, our data from Optimal Blue is consistent with these facts—the mean rebate across rate sheets for our entire sample is approximately 1, with a standard deviation of 2.2.

To address the problem of time-varying sets of note rates, we derive a constant-rebate note rate, which we argue provides a reasonable index of mortgage rates for our analysis. Our main analysis uses “Rate101,” the note rate that yields a rebate of +1 point (or  $p_{YSP} = 101$ ), which is anecdotally a typical rebate that borrowers choose. We construct Rate101 for each rate sheet by interpolating between different offers and then use the daily median value across our sample of rate sheets as our baseline.<sup>14</sup>

---

<sup>13</sup>“BofA to exit correspondent mortgage business,” Reuters, Aug 31, 2011.

<sup>14</sup>The interpolation is required because, as Figure 2 shows, rebates are quoted in decimals but note rates are generally quoted in 1/8 of a percentage point. In the rare case that the lender only offers rebates above or below 101, we exclude the lender from our sample.

The top panel of Figure 3 shows the evolution of Rate101 over our sample. Rate101 started above 6 percent in 2008, prior to the monetary policy actions that started in late November of that year, and reached its low point near the end of 2012 at close to 3 percent. The bottom panel shows how Rate101 compares to two alternative choices, Rate100 (for a rebate of 0) and Rate102 (for a rebate of +2), and also to the widely quoted Freddie Mac Primary Mortgage Market Survey rate.<sup>15</sup> The figure shows that over 2010–2014, the rate change per 1 point in rebate was quite stable around 20 basis points, and that the Freddie Mac rate was on average quite close to our Rate101 (thereby validating our choice of +1 as a “typical” rebate). Earlier in our sample, however, the effect on the note rate of a change in rebate was much larger: at times, a borrower could get a 100 basis point reduction in his or her mortgage rate by accepting a 1-point lower rebate.<sup>16</sup> This in turn made it attractive to borrowers to get relatively lower rebates (i.e. 0 instead of +1); this is reflected in the Freddie Mac survey rate, which is closer to Rate100 over this early period.

Rate101 changes from one day to the next, but to study daily pass-through we will want to hold the note rate constant over time. Thus, we introduce the notation  $p_{YSP}^{Rate101,t}(s)$ , which means the time- $s$  primary market price of a loan with a note rate of time- $t$  Rate101. By definition,  $p_{YSP}^{Rate101,t}(t) = 101$ . We define  $p_{TBA}^{Rate101,t}(s)$  similarly.

### 3.2 Measuring the Value of a Mortgage in the Secondary Market

The main secondary market for mortgages is called the To-Be-Announced or TBA market. The TBA market is a forward market for mortgage-backed securities (MBS) guaranteed by Fannie Mae, Freddie Mac, and Ginnie Mae, and is one of the largest fixed income markets in the world (Vickery and Wright, 2013). In this market, buyers (such as mutual funds, pension funds, foreign investors, and, in recent years, the Federal Reserve) promise to buy a pool of mortgages meeting certain specified characteristics at a specific date in the future. The majority of newly produced agency MBS pools are traded in this market.<sup>17</sup> Agency MBS are typically only traded in coupon increments of 50 basis points (bps).

As mentioned earlier, the guaranteeing agency, in our case Fannie Mae, charges an ongoing insurance premium or guarantee fee (“g-fee”) as well as an upfront premium depending on loan characteristics, known as “loan-level price adjustment” (LLPA). The exact values of

---

<sup>15</sup>The Freddie Mac rate is based on a weekly survey of about 125 lenders, asking them for their “most popular” rate and rebate combination on first-lien prime conventional conforming home purchase mortgages with a loan-to-value ratio of 80 percent. We focus on the series for 30-year fixed-rate mortgages.

<sup>16</sup>This was due to a narrowing of MBS price differences across coupons, as also discussed in Fuster and Willen (2010).

<sup>17</sup>The remainder are pools that either do not meet the criteria for the TBA market or are so-called spec pools, meaning groups of loans with certain characteristics that are desirable to investors (for example, an investor might want low credit score loans because of their more attractive prepayment properties).

g-fees are not disclosed, but we know that they vary over our sample. Up to 2012, the typical g-fee was roughly 22 bps, meaning that Fannie charged an annualized fee of 0.22 percent of the unpaid principal balance of a loan every month.<sup>18</sup> In 2012, FHFA, Fannie and Freddie’s regulator, announced two 10 basis point increases. The first, on January 1, raised the g-fee to 32 bps and the second, on September 1, raised it to 42 bps, where it remained for the rest of our sample. The upfront LLPA for the base-case loan characteristics only changes once over our sample period, from 25 bps prior to December 23, 2010, to 50 bps.<sup>19</sup>

In our baseline calculation for  $\phi$ , we assume that the value of a mortgage with note rate  $Rate_{101}$  in the secondary market is given by the interpolated TBA market price of an MBS with coupon  $Rate_{101} - gfee$ , subtracting the upfront LLPA. For instance, if  $Rate_{101} = 4.56$ ,  $gfee = 0.32$ , and  $LLPA = 0.25$ , then our the secondary market value equals:

$$p_{TBA}^{Rate_{101}}(s) = p_{TBA}^4(s) + (24/50) * (p_{TBA}^{4.5}(s) - p_{TBA}^4(s)) - 0.25,$$

that is, we interpolate between the prices of the 4.0 and 4.5 percent coupons.<sup>20</sup>

### 3.3 Data Sources

The MBS price data is from J.P. Morgan and, for simplicity, we assume that all loans are sold to Fannie Mae. For our baseline analysis, we normalize all prices to be 45-days-to-settlement using a weighted combination of the 1-, 2-, and 3-months-out contracts.<sup>21</sup> This is the relevant metric if the time between rate lock (prior to loan origination) and delivery in a TBA trade is 45 days, which is close to the approximate average lag from application to origination. Different normalizations do not materially affect our results since the gaps between the prices are generally stable.

To measure time-series variation in demand, we use new loan applications from the confidential version of the Home Mortgage Disclosure Act (HMDA) dataset, which captures a large share (roughly 90 percent) of mortgage applications and originations. The confidential

---

<sup>18</sup>The value of 22 bps is obtained from Fannie Mae 10Ks; we take an average of the average guarantee fee over 2005–2008. After 2008, LLPAs affect the level of guarantee fees reported in the 10Ks, since they are not separated out.

<sup>19</sup>This value includes the 25 bps “adverse market delivery charge” that the GSEs started charging in March 2008; see <https://www.fanniemae.com/content/announcement/0721.pdf>. Fannie Mae’s current LLPA matrix is available at <https://www.fanniemae.com/content/pricing/llpa-matrix.pdf>.

<sup>20</sup>Our calculation above implicitly presumes that a loan always gets securitized in a coupon below  $Rate_{101} - gfee$ , whereas in reality originators have the option to securitize a loan in any coupon rate at or below  $Rate_{101} - 0.25$ . Fuster et al. (2013) provide a detailed discussion of the decision into which coupon to pool; what is important for us here is that the differences across these options are typically small.

<sup>21</sup>Depending on the time of the month, 45 days falls in between the 1- and 2-month contracts or the 2- and 3-month contracts and we interpolate accordingly to avoid spurious jumps in MBS prices. For details about the timing, see [http://www.mortgagenewsdaily.com/mortgage\\_rates/blog/162276.aspx](http://www.mortgagenewsdaily.com/mortgage_rates/blog/162276.aspx).



version of the data contains exact application and action (i.e. accept or reject) dates, which allows us to count applications on a daily level. We include all first-lien, single-family loans in our measure, including applications for refinancing and purchase loans. The available data covers 2008 through October 2014. In our lower-frequency analysis in Section 6, we use total monthly applications, normalized by the number of business days in a given month.

We also use the HMDA data to measure lender concentration, by calculating the market share of the top four lenders (following [Scharfstein and Sunderam, 2013](#)) at a monthly frequency. This measure is specific to the MSA in question, which is Los Angeles in our baseline case (since we are looking at rate/rebate offers for Los Angeles), although the overall trends in market concentration are similar across the country during the period we study.

As a measure of labor costs, we use real estate credit payrolls, divided by real estate credit employment. Both series are compiled by the Bureau of Labor Statistics and available at a monthly frequency. To measure the volatility of interest rates, which may affect originators' hedging costs, we use the monthly average Merrill Lynch Option Volatility Expectations (MOVE) index, an index of the normalized implied volatility on 1-month Treasury options, weighted on the 2-, 5-, 10-, and 30-year contracts.

Finally, as discussed in Section 2.3.1, for our alternative measure of intermediary margins  $\pi$  we use servicing valuation multiples provided to us by the Mortgage Industry Advisory Corporation (MIAC). We have monthly data from the start of our data in October 2008 until September 2014; we assume that the value in October 2014 is the same as in September 2014.<sup>22</sup> For the valuation of the remaining interest strip (after g-fee and 25 bps servicing are deducted from the coupon), we interpolate between the TBA prices of the surrounding coupons, as we do in our baseline calculation.

## 4 Event Studies: QE Announcements

In this section, we use the concepts derived above to explore the effect of announcements about the Federal Reserve's large-scale asset purchase program (and one other unconventional monetary policy measure, date-based forward guidance) over the 2008 to 2013 period. We do this first because the results are of inherent interest: an explicit purpose of asset purchases (or quantitative easing, QE) was to drive down the cost of mortgage credit for consumers. But, more broadly, the monetary policy announcements provide a sort of laboratory to understand the transmission of shocks in the secondary mortgage market to the

---

<sup>22</sup>An alternative multiple series, available at <https://www.federalreserve.gov/publications/2016-capital-rules-mortgage-servicing-assets-Evolution-of-the-Mortgage-Servicing-Market-since-1998.htm> (Figure 6), produces almost identical results.

primary market. The six events we focus on are:

1. **“QE1”**: November 25, 2008. Announcement by the Federal Open Market Committee (FOMC) that the Fed would purchase up to \$600B in agency MBS and agency debt.
2. **QE1 Expansion**: March 18, 2009. The FOMC announced the program would be expanded: purchases of agency MBS and agency debt increased by \$750B; purchases of Treasury securities increased by \$300B.
3. **Forward Guidance**: August 9, 2011. The FOMC changed its statement language to “exceptionally low levels for the federal funds rate at least through mid-2013” from “exceptionally low levels for the federal funds rate for an extended period.”
4. **“QE3”**: September 13, 2012. The FOMC announced an open-ended commitment to purchase \$40B agency MBS per month until labor market improved “substantially.”
5. **Taper Tantrum**: June 19, 2013. Fed Chairman Bernanke publicly discussed “tapering” asset purchases: “If the incoming data are broadly consistent with this forecast, the committee currently anticipates that it would be appropriate to moderate the monthly pace of purchases later this year.”
6. **Non Taper**: September 18, 2013. Contrary to market expectations, the FOMC opted *not* to reduce its MBS purchases.

Our choice of events is, by its nature, somewhat arbitrary but our goal was to focus on news events that surprised the markets and led to significant movements in the MBS market, as all of these did.<sup>23</sup>

To explore the evolution of prices in the primary and secondary mortgage markets, we start with the Rate101 for date  $\bar{t}$ . We calculate  $p_{TBA}^{Rate101,\bar{t}}(\bar{t})$  for that day, while  $p_{YSP}^{Rate101,s}(\bar{t}) = 101$  by definition. We then consider different dates  $t$  and calculate  $p_{TBA}^{Rate101,\bar{t}}(t)$  and  $p_{YSP}^{Rate101,\bar{t}}(t)$ , which no longer necessarily equals 101. To illustrate this, consider the announcement of QE3 in September 2012 (the bottom left panel of Figure 4). Table 2 shows how we compute the two series for the date immediately prior to the announcement on September 12 and the date immediately following, September 13, illustrating that we keep the note rate constant over time and allow the primary and secondary market prices to move. The two series imply that 55 percent (0.6/1.1) of the secondary market price increase was passed through to borrowers

---

<sup>23</sup>To enter our sample, the date had to be among either the 25 days in our sample with the biggest increases or among the 25 days with the biggest decreases in the yield on the current coupon MBS (from J.P. Morgan) *and* had to coincide with a major announcement regarding Federal Reserve policy. As an example of a date that did *not* meet our criteria, the announcement of “QE2” on November 3, 2010 had little effect on bond markets because anticipation of QE2 had led to a substantial rally over the previous three months.

	$\bar{t} = 9/12/2012$	$t = 9/13/2012$	
$Rate101, \bar{t}$	3.47	3.47	
$p_{YSP}^{Rate101, \bar{t}}$	101	101.6	$\Delta p_{YSP} = 0.6$
$p_{TBA}^{Rate101, \bar{t}}$	102.7	103.8	$\Delta p_{TBA} = 1.1$

Table 2: Calculation of Pass-Through for the Announcement of QE3 on September 13, 2012

on the day of the QE3 announcement. The figure does the same calculation for subsequent dates and shows that  $p_{TBA}$  drifted down over the next few days but  $p_{YSP}$  stayed more or less the same so the effective pass-through was somewhat higher at longer horizons.

Looking across the six events in Figure 4 now, several features are worth noting.<sup>24</sup> First, the high frequency relationship between  $p_{TBA}$  and  $p_{YSP}$  is quite close. Almost everywhere, changes in the secondary market price lead to changes in the primary market price on the same day. It is important to stress that there is nothing mechanical about this relationship in the data as the two time series come from completely different data sources:  $p_{TBA}$  is generated from global financial markets data and  $p_{YSP}$  comes from offers to loan officers working with individual borrowers.

The second notable feature of Figure 4 is that the pass-through appears to vary significantly across the different events. For example, borrowers received less than 50 percent of the effects of the QE1 Expansion announcement in March 2009. In contrast, the effects of the initial announcement of QE1 in November 2008, the Taper Tantrum in June 2013, and the Non-Taper in September 2013 were passed on to borrowers almost entirely. What factors can explain these differences? In Section 2.1 above, we argued that the degree of pass-through depends on the shape of the supply curve. Recall that an increase in prices in the secondary market, all else equal, leads to an increase in the demand for intermediation and that a higher slope of the cost curve implies lower pass-through, as depicted in Figure 1.

To explore the effect of demand on pass-through, we add the daily volume of new loan applications to each figure. As noted already, the pass-through was very high for the initial QE1 announcement but much lower for the QE1 Expansion four months later. Consistent with the idea that application volume matters, we see a large difference in volume around the announcements. Indeed, at 35,000 applications per day, the volume on November 24, 2008 was at a historically low level, so there was likely some capacity slack that allowed for pass-through of increased MBS prices to borrowers. In contrast, the day before the announcement of the QE1 Expansion, there were 60,000 applications and that number spiked to over 100,000

<sup>24</sup>For the first two events, we use Rate100 instead of Rate101 because Rate100 was more representative of the note rates borrowers were choosing at that time (see the earlier discussion of Figure 3).

following the announcement; consistent with the idea of intermediaries facing convex costs in quantity, the pass-through was lower. The last two events are also instructive. In June 2013, bond prices fell dramatically and, not surprisingly, demand did not increase and the pass-through was high. In September 2013, bond prices rose but from a low level induced by the original Taper Tantrum, meaning that even after the increase, refinancing demand remained modest.

A third feature of Figure 4 is that the price of intermediation  $\phi$  appears to rise over time. If we compare the original announcement of QE1 with the Non-Taper, the pass-through was higher in the latter, as already noted, but the gap between intermediaries' revenue ( $p_{TBA}$ ) and what they pay the borrower for the loan ( $p_{YSP}$ ) is much higher in the latter. This increase in  $\phi$  could, according to the theory, result from higher volume, but the data appears to reject that — application volume in September 2013 was very similar to application volume before the QE1 announcement, but  $\phi$  was much higher.

In the next two sections, we will consider questions raised here using an econometric approach. How big is pass-through? Does volume systematically affect pass-through? Is pass-through higher for decreases in prices versus increases? Did the price of intermediation increase over time? What accounts for that increase?

## 5 High-Frequency Pass-Through

We now study to what extent changes in MBS prices from one day to the next are reflected in the rebates offered on rate sheets. To do so, we follow the method explained when discussing Table 2 to calculate one-day changes in  $p_{YSP}$  and  $p_{TBA}$  for each day in our sample. The top panel of Figure 5 shows a scatter plot of the two sets of changes. From the figure, it is clear that there is a very strong relationship between price changes in the primary and secondary market, though the pass-through appears slightly stronger for negative changes than positive changes.

Table 3 shows the results from a linear regression performed to quantify the amount of pass-through in Figure 5. Column (1) shows that the average pass-through coefficient is 0.92, and that 88 percent of the variation in the rebate changes is explained by the variation in MBS price changes. Pass-through of price decreases appears to be stronger than the pass-through of price increases (column (2)). Notably, the asymmetry of pass-through is statistically and economically significant, with price decreases being passed through approximately 100 percent within this one-day window. Price increases, in contrast, are only passed through about 80 percent on the first day.<sup>25</sup> In column (3), we add the price change from the previous

---

<sup>25</sup>Figure 5 suggests potential non-linearity in the relationship between secondary and primary price

day, and find that there is some additional pass-through of lagged positive (but not negative) changes, so that over two days, approximately 87 percent of MBS price increases are reflected in primary market rebates. (In untabulated results, we find that the effects for further lagged MBS price changes are statistically indistinguishable from zero.)

In the remaining two columns we explore a potential driver of time variation in pass-through, namely the level of demand, following the discussion in Section 2. Specifically, we interact the changes in the MBS value with the standardized level of new applications from HMDA (lagged by one day), so that the coefficient on the interaction term corresponds to the effect on daily pass-through of a one-standard deviation increase in application volume. Column (4) shows that indeed, a higher level of applications lowers the pass-through of MBS price increases to rebates. For instance, the coefficients imply that if applications are two standard deviations above average, only 59 percent ( $= 0.781 - 2 * 0.096$ ) of a price increase is passed through. The final column shows that this interaction effect remains unchanged if we add the lagged price change as well. The uninteracted lagged application level is not significantly related to changes in  $p_{YSP}$ .

The bottom panel of Figure 5 graphically illustrates how differences in application volume at a lower frequency are related to time variation in the pass-through coefficients. The figure plots the coefficients from regression (2) shown in Table 3, but estimated separately year-by-year (except that we pool the 2.5 months of data from 2008 with the 2009 data). The figure shows that for MBS price increases, the estimated pass-through was lowest in 2012, when applications were at their highest levels. The pass-through then increased over the following two years, as application volumes dropped; in 2014, the year with the lowest mortgage demand in our sample, the pass-through of positive price changes exceeds 0.9. The pass-through of price decreases (the upper line) is fairly stable and close to 1, except in 2013 when it goes up to 1.1.

In Appendix B we discuss an alternative methodology where changes in mortgage *rates* are regressed on changes in MBS *yields* (on the so-called current coupon, which is the market benchmark), at the daily, weekly, or monthly frequency. With Rate101, we obtain a daily pass-through similar to what we found in this section. Perhaps surprisingly, the pass-through coefficients exhibit a slight decrease when we look at longer horizons (especially monthly); the reasons for this will be further discussed in the next two sections. Finally, we find much stronger pass-through using our Rate101 rather than the commonly used Freddie Mac rate, suggesting that there is less (non-random) measurement error in our rate measure.

---

changes. In untabulated regressions, we find that adding quadratic MBS price changes results in a significant negative (but relatively small) coefficient for price declines, and an insignificant positive coefficient for price increases. None of the other results in this section are qualitatively altered by adding squared price changes to the regressions.

The take-away from these regressions is that at high frequencies, prices in the primary mortgage market strongly respond to the secondary market. In a sense, this should not be surprising, given that lenders issue new rate sheets every day and sometimes more than once a day. But this result does contrast with findings of infrequent price adjustments in other consumer finance markets, where rates offered to consumers only react slowly to changes in monetary policy and market rates (e.g. [Ausubel, 1991](#); [Driscoll and Judson, 2013](#)). One possible explanation for the difference is that brokers and correspondent lenders, who are doing the “shopping” for mortgage borrowers, make the market more competitive than other consumer finance markets. [Allen, Clark, and Houde \(2014\)](#), in their study of the Canadian mortgage market, argue that brokers play such a role there.

Although these high-frequency regressions demonstrate a direct and strong link from MBS prices to the primary market, the result that increases in MBS prices result in incomplete pass-through leaves open the possibility of lower frequency movements in the price of intermediation, which we turn to next.

## 6 Evolution of the Price of Intermediation Over Time

Figure 6 shows the evolution of our baseline estimate of the price of intermediation at the monthly frequency (a daily version is shown in Appendix Figure A.1). This cost is calculated as the secondary market value of the Rate101 loan  $p_{TBA}^{Rate101}$  (net of payments to Fannie Mae) less the primary market value  $p_{YSP}^{Rate101}$  which is, by definition, 101:

$$\phi(t) = p_{TBA}^{Rate101,t}(t) - p_{YSP}^{Rate101,t}(t) = p_{TBA}^{Rate101,t}(t) - 101. \quad (6)$$

We estimate that over the whole sample period, the price of intermediation averaged 142 basis points. The standard deviation was large (61 basis points), as evidenced by the fact that our estimated values range from 0 to 300 basis points.

In Section 2.3.1, we raised the issue of how the costs and benefits of servicing might affect the value of a mortgage in the secondary market. To address this, we calculate  $\pi$ , an alternative measure of intermediation margins where the 25 basis points of servicing cash flow are valued using a separate multiplier from an industry source (MIAC) that takes into account other costs/benefits of servicing. Figure 6 shows that the time series of  $\pi$  displays a somewhat less pronounced upward trend than  $\phi$ . In fact, early in our sample period,  $\pi$  is larger than  $\phi$ , meaning that the MIAC servicing multiples were higher than the MBS-price-implied multiples, which in turn means that the additional benefits from servicing exceeded the costs. Anecdotally, this was typical in the pre-crisis period more generally. However,

in late 2009 the relationship between  $\pi$  and  $\phi$  flips, meaning that the additional costs of servicing exceeded the benefits. This implies that the present value of net servicing costs increased over our sample period, something we will address later.

The general patterns displayed in our series, large high frequency changes and an upward trend over time, are consistent with other measures of the price of intermediation. For example, the mortgage gain-on-sale margins reported by large banks show similar patterns (see Appendix C). Moreover, in Appendix D we present a longer time series of an approximated  $\phi$  based on the Freddie Mac PMMS rate, which shows that the level of  $\phi$  in late 2008 was not unusually low relative to the previous years.

## 6.1 Determinants of Variation in the Price of Intermediation

We now turn to potential drivers of the variation in the price of intermediation. Figure 6 plots the time series of four key variables that proxy for either the costs of intermediaries or their ability to earn excess profits: (1) new application counts taken from the HMDA data; (2) the payroll per employee in the real estate credit sector, obtained from the BLS; (3) interest rate volatility as measured by the Merrill Lynch MOVE index; (4) market concentration, as measured by the market share of the top four lenders (HMDA).

Applications proxy for increasing marginal costs, or increased pricing power due to limited origination capacity, as discussed earlier. Indeed, panel B of Figure 6 shows a strong positive correlation between  $\phi$  and application volume, in particular at high frequencies (for example, the two spikes in application volume in 2008 and early 2009 correspond to spikes in  $\phi$  almost exactly). Employee payroll is an important component of origination costs and has seen an upward trend over our sample period (panel C). Two other potential drivers of higher intermediation prices are interest rate volatility and market concentration. Higher rate volatility increases hedging costs and concentration increases intermediary market power. Panels D and E show that both series were *decreasing* for much of our sample and, further, high frequency changes do not appear to coincide with changes in  $\phi$ .

To analyze these relationships formally, we begin by regressing  $\phi$  on these variables, at the monthly frequency, both in levels and changes. Furthermore, we add Rate101 as an additional control variable to ensure that the correlations we document are not driven by the downward trend in rates over the time period we study. A main reason why we run the levels regressions (which could pick up spurious relationships between trending variables) is that these allow us to add a linear time trend, so we can test whether there was an increase in  $\phi$  that is not accounted for by other explanatory variables.

Since we use a measure of the price of intermediation as the dependent variable, and

applications (quantity) on the right-hand side, one concern is that the coefficient on applications is biased or the correlation is spurious (since quantity is of course endogenous to price). The first thing to keep in mind is that a higher price should correlate with a reduced quantity (holding everything else fixed), so this effect should bias our coefficient of interest downward or even lead to a negative coefficient, while we will find a strongly *positive* coefficient. Thus, this endogeneity concern works against our main finding.

However, to further assess how important this bias could be, we also employ an instrumental variables (IV) strategy. As our instrument, we need something that affects the application volume without directly affecting the price of intermediation. We use a proxy for borrowers' incentives to refinance that is independent of the current price of intermediation, namely the difference between the weighted average coupon (WAC) on outstanding conventional MBS and the ten-year constant maturity Treasury yield. This measure is closely correlated with borrowers' rate incentive to refinance (which would be given by the WAC minus the current mortgage rates) but is not affected by current developments in the mortgage market itself. In terms of the supply/demand relationship shown in Figure 1, one can think of an increase in (WAC–Treasury yield) as proxying for an upward shift of the demand curve, which allows us to trace out the effect on the equilibrium price.

Panel A of Figure 7 shows the evolution of the WAC on outstanding conventional mortgages and the ten-year Treasury yield. Not surprisingly, the latter is responsible for most of the high-frequency variation in the difference between the two series. Panel B shows that this difference is closely correlated with application volume over our sample period, both in levels ( $\rho = 0.87$ ) and in monthly changes ( $\rho = 0.73$ ).

Turning to the regressions, Tables 4 and 5 show that an increase in new application volume is strongly associated with a higher price of intermediation. This result holds in levels and changes, with and without other explanatory variables, and whether or not we instrument for applications as described above. In these regressions, all explanatory variables except Rate101 are normalized to have a mean of zero and a standard deviation of one.

Focusing first on the OLS results in levels, we see a marginally significant correlation between raw  $\phi$  and applications (column 1 of Table 4). Adding the linear time trend substantially increases the coefficient on applications in column (2), and the time trend and applications jointly explain 84 percent of the variation in  $\phi$  (while the time trend alone would explain 52 percent). The tight fit is illustrated by the chart below the table. In terms of the magnitude of the effect, a one-standard deviation increase in applications is associated with a 36 basis point increase in  $\phi$ . Introducing the additional control variables in column (3) only slightly lowers the coefficient on applications; adding Rate101 slightly reduces it further, even though the coefficient on Rate101 itself is not significant.



The IV versions of the same regressions lead to a slightly higher coefficient on applications, as one would expect if higher intermediation prices tend to lead to fewer applications, other things remaining equal. The first-stage F-statistics, shown in the last row of the table, indicate a strong explanatory power of our excluded instrument for applications, the potentially endogenous variable. Across specifications, the positive time trend in  $\phi$  that is not explained by the other explanatory variables is estimated to be around 1.6–2.6 basis points per month, corresponding to 20–30 basis points per year.

The regressions in changes, shown in Table 5, generate similar coefficients on applications. One thing to note with the IV version is that the change in Rate101 is strongly correlated with the change in the instrument, which leads the coefficient on applications to strongly increase between columns (5) and (6). As a consequence, it is preferable to not include Rate101 as a regressor in the IV version.

Table 6 separately analyzes the effects of applications and the other controls on the two components of  $\phi$ , the origination costs ( $\pi$ ) and the present value (PV) of servicing costs ( $\phi - \pi$ ), following the decomposition discussed in Section 2.3.1. We show the IV version of this analysis only; the conclusions are similar when not instrumenting for applications. The table shows that applications are very strongly and positively correlated with  $\pi$ , whether the estimation is done in levels or changes; this finding is consistent with origination costs or profits increasing with demand. In contrast, the PV of servicing costs displays a negative correlation with applications, and the effect is statistically significant in the estimation using monthly changes. This finding is likely driven by the fact that mortgages originated at a time of high application volume tend to have a shorter expected life span (precisely because  $\phi$ , or alternatively the gap between the mortgage note rate and MBS yields, is high at those times, and expected to decrease in the future).<sup>26</sup>

Interest rate volatility is positively related to origination costs, as would be expected given that the cost of hedging increases when volatility is high. In contrast, volatility is significantly and negatively related to  $\phi - \pi$ . The explanation is similar to the one offered for applications: when rate volatility is high, this decreases the expected life span of a mortgage (since the borrowers’ prepayment option becomes more likely to be in-the-money going forward) and thus the expected PV of servicing costs decreases. These two offsetting effects explain why overall,  $\phi$  is not significantly related to volatility.

The measure of payroll per employee in the real estate credit sector, which trended upward over the sample period, is positively related to  $\pi$ , suggesting that an increase in the

---

<sup>26</sup>Indeed, in unreported results, we find that adding the “primary-secondary spread” as a regressor brings the coefficient on applications in the  $\phi - \pi$  regression much closer to zero and eliminates its statistical significance.

labor cost involved in the origination process could be an important part of the upward trend in the price of intermediation. In monthly changes, payroll is not significantly related to  $\pi$ , but month-to-month changes likely contain significant noise. Surprisingly, the correlation between payroll and the PV of servicing costs is slightly negative, though this correlation is only statistically significant in levels.

There is no evidence that changes in market concentration affected either component of intermediation prices over this period—which is not surprising, given that concentration trended down over this time (so regressing  $\phi$  or  $\pi$  on concentration alone actually yields a negative and significant coefficient). The time trend in the levels version of the regressions is stronger and more statistically significant for the PV of servicing costs, suggesting that this component plays an important role in the overall increase of  $\phi$  (further discussed below).

Figure A.3 in the Appendix shows that our measure of  $\phi$  evolves similarly if we change the assumed note rate (to Rate100 or Rate102), the assumed FICO score (to 680 instead of 750), the MSA (New York instead of Los Angeles), or the loan size (\$150,000 instead of \$300,000). As a consequence, regressions with the  $\phi$ ,  $\pi$ , or  $\phi - \pi$  series based on these alternative assumptions yield results very similar to those shown in Tables 4 to 6. These findings provide robustness checks with respect to our assumptions on loan characteristics, but below we will also discuss what they mean for the interpretation of the results.

## 6.2 Interpretation

The two key findings of our regression analysis of low frequency movements in  $\phi$  and  $\pi$  are the correlation with application volumes and the time trend. We discuss each in turn. We view the strong positive correlation between  $\phi$  and application volume documented above as indicative of limited capacity in mortgage intermediation. To the extent that  $\phi$  (or  $\pi$ ) includes profits (or high rents for the factors of production), why does capacity not expand so that profits get competed away? Potential explanations include reluctance among lenders to add additional capacity if they think the increase in volume is relatively short-lived, as is often the case with refinancing booms. In addition, new entry into the mortgage intermediation business is costly; for instance, an intermediary has to fulfill a number of requirements (such as a minimum net worth) to be able to securitize mortgages through the GSEs.

To further support our interpretation, we present evidence based on the processing times for mortgage applications according to the HMDA data. Figure 8 shows, at the monthly frequency, the number of new mortgage applications plotted against the median processing time of applications submitted in that month (i.e. the number of days between the application date and the “action” date, when either the loan closes, or the application is denied or

withdrawn). The figure shows that when the number of new mortgage applications increases, it takes longer to process additional applications. The correlation between the two series in levels is 0.74; in changes it is 0.63. The variation in processing times can be substantial: for instance, it goes from 25 days to about 40 days following the November 2008 Fed announcement. This evidence suggests that the capacity of mortgage originators to process applications is limited.

A regression analysis, shown in Table 7, confirms the significance of this relationship: a one-standard deviation increase in monthly applications is associated with an increase in processing time of three to four days. It also shows that conditional on application volume, there was a significant upward trend in processing time over this period.<sup>27</sup> This finding is consistent with industry surveys indicating that the labor intensity of underwriting has increased substantially since 2008; see Appendix C for further discussion.

This leads to the question of how to interpret the positive time trend in  $\phi$ . An important component of this trend comes from the relative decrease in the effective value of servicing rights based on the MIAC multiples relative to their MBS-price-implied value—which is what is driving the difference between  $\phi$  and  $\pi$ . In other words, net servicing costs have increased over 2008–2014. This result is consistent with evidence of an increase in the direct costs of servicing over this period (see Appendix C) and also the increased regulatory cost for banks of holding mortgage servicing rights under Basel III (see Hendricks et al. 2016 for a thorough discussion). Consistent with the latter channel, there has been a significant shift of servicing activity from banks to non-banks since 2010.<sup>28</sup>

But even after accounting for the effect of servicing valuations, there remains a positive time trend in  $\pi$  as long as we do not control for payroll (see column (2) of Table 6). This could be due to the increased labor intensity of underwriting noted above, or the need for more qualified people doing it—an explanation consistent with the positive correlation with payroll costs (though an alternative interpretation could be that employees in the real estate credit sector simply earn higher rents when the profitability of originations increases).

There are a number of potential explanations that seem less consistent with our evidence. One such explanation is that the positive relationship between  $\phi$  ( $\pi$ ) and applications, and/or the positive trend in the series, are due to changes in the expected life span of the originated loans. If there is a constant per-period cost of an open loan (e.g. due to the liability/putback

---

<sup>27</sup>Calendar months are also important in explaining variation in processing time; in particular, processing times (for given application volume) tend to increase in December. The last column in the table shows that once we add indicators for December and January to the regression in changes, the adjusted  $R^2$  increases dramatically and the constant becomes significant (indicative of the positive time trend).

<sup>28</sup>According to Kaul and Goodman (2016), the share of single-family mortgages (in terms of remaining balances) serviced by non-banks increased from 6 percent in 2010 to 31 percent in 2015.

risk, discussed below), variation in the expected life span would translate into differences in  $\pi$  (as originators want to be compensated upfront if expected lifetime costs increase).

We already noted above that the PV of servicing costs,  $\phi - \pi$ , is significantly related to changes in the loan's expected life due to changes in interest rate volatility, and perhaps also due to changes in rates. However, the strong correlation between origination costs  $\pi$  and applications is not affected by adding these variables to the regression (if anything, the opposite is the case). We have also collected direct proxies for the expected lifespan of newly originated mortgages and find that they are not positively related to  $\phi$ —see Appendix E.

Other evidence also does not point toward an important role for anticipated per-period costs (other than those already captured in the servicing valuation) in the increase in  $\phi$ . One such cost is related to potential legal liabilities. In the wake of the crisis, the GSEs have aggressively enforced “representations and warranties,” forcing many entities that had sold them mortgages to buy back (mostly delinquent) mortgages where some flaw in the underwriting or documentation was found. This “putback risk” has commonly been cited as a reason behind the tighter lending standards in recent years, since lenders want to avoid incurring this risk on new loans going forward.<sup>29</sup>

If increased expected putback costs were priced directly by intermediaries, we would expect that  $\phi$  would have increased more on some types of loans than others. In particular, this theory would suggest that riskier loans should have seen a larger increase in  $\phi$  (since they are more likely to lead to a delinquency followed by a repurchase and a credit loss).<sup>30</sup> Furthermore, loans with a longer expected life span should have seen a larger increase in  $\phi$ , since they will remain a potential liability for longer. However, as shown in Figure A.3,  $\phi$  did not increase more for FICO 680 loans than for FICO 750 (if anything, the contrary), even though the former have higher delinquency risk. Also, there is no evidence that smaller loan sizes (150k versus 300k), which are associated with slower prepayment speeds, command a higher  $\phi$  for the same note rate.<sup>31</sup> There is some divergence across  $\phi$  for note rates Rate102 versus Rate100 in 2012, with  $\phi$  being higher for lower note rates (with longer expected life).

---

<sup>29</sup>For instance, the Federal Reserve's Senior Loan Officer Opinion Survey in April 2013 found that “Three-fourths of banks also cited the risk of putback of delinquent mortgages by the GSEs as an important factor restraining their current ability or willingness to approve home-purchase loans, and . . . a large fraction of banks reported an increase in the importance of this factor over the past year.” See <http://www.federalreserve.gov/boarddocs/snloansurvey/201305/default.htm>.

<sup>30</sup>Goodman and Zhu (2013) show that the incidence of putbacks is substantially higher for low-FICO loans, and also provide additional discussion of putback patterns in recent years.

<sup>31</sup>It is somewhat puzzling that  $\phi$  does not vary much with loan size, given that we think of origination costs as having a substantial fixed component (and thus might expect a larger proportional charge for smaller loan sizes). However, we have verified on lender rate sheets that most of them have at most minor adjustments for small loan sizes. It is still possible that borrowers with small loans will typically pay proportionally more upfront (or take a higher note rate) to cover the fixed costs of loan origination not captured in our measure.

But the difference remains relatively small compared to the overall increase in  $\phi$  over the sample period, and reverts back to zero in 2013/early 2014. That said, it is plausible that the increased aversion to putbacks was a significant driver of the upward trend in origination costs (for instance, due to the need for more qualified underwriting personnel, and more time spent on processing each loan, as discussed above).

Finally, as noted above, our measure of  $\phi$  is almost unchanged if we use New York instead of Los Angeles as the MSA in our rate sheet search. This is the case even though market concentration (as measured in the HMDA data) was somewhat higher in Los Angeles over 2008–2009 but then fell quite steeply and has been lower than in New York since 2012. This again means that we do not find evidence suggesting that market concentration (or changes therein) explains the pattern in intermediation costs that we document.<sup>32</sup> However, our findings are not necessarily inconsistent with those of [Scharfstein and Sunderam \(2013\)](#) — they simply mean that at least over the period since 2008, the effects of market concentration did not occur at the level we study (i.e. intermediary rate sheets) but could instead occur through splitting the rebate between the loan officer and borrower.

## 7 Economic Implications

In this section, we first use our estimates of the price of intermediation to measure how much households spent on intermediation during the period we study and conduct several counterfactual exercises to decompose household expenditures. We then translate these effects into equivalent effects on mortgage rates and clarify the relationship between our measure of the price of intermediation and a commonly used alternative, the spread between primary market rates and secondary market yields.

### 7.1 Expenditure on Intermediation

Over the 73-month period we cover in the paper, U.S. households refinanced 6.4 trillion dollars of mortgages and used 3.5 trillion dollars of new mortgage debt to purchase homes.<sup>33</sup>

---

<sup>32</sup>Similarly, at the national level, concentration in mortgage lending has been decreasing over our sample period. According to [Inside Mortgage Finance \(2016\)](#), the overall market share of the top 10 lenders in the United States overall was 75.7 percent in 2010 before declining to 63 percent in 2012 and 45.4 percent in 2014. Also, the weighted-average county-level “Top 4” share shown in Figure 1 of [Scharfstein and Sunderam \(2013\)](#) fell from over 0.38 in 2010 to below 0.3 in 2013 and 2014, and is thus close to its longer-term average for the period since 1994. Concentration in servicing has likewise been decreasing, with the market share of the top 10 servicers falling from 66.5 percent in 2008 to 54.8 percent in 2014.

<sup>33</sup>These figures are based on HMDA data, and may thus slightly understate the actual total, since about 10 percent of loans are not covered in the HMDA data.

According to our estimates, households implicitly paid 147 billion dollars to financial intermediaries for their services in these transactions (95.7 billion dollars for the refinances and 51.8 billion dollars for the purchase mortgages), or about 25 billion per year on average. We arrive at this estimate by using our monthly average estimates of  $\phi$  from Section 6.1 and multiplying them by the dollar amount of monthly originations.<sup>34</sup> One obvious question is how large a number this is. As a point of comparison, Hurst et al. (2016) find that the GSE policy of maintaining uniform mortgage rates across locations (rather than letting rates vary with credit risk) implied a redistribution across U.S. regions of 14.5 billion dollars (in NPV terms) over the 2007–2009 period, or roughly 5 billion per year.

The top panel of Figure 9 shows our estimate of the time series of monthly expenditures, which range from less than 1 billion dollars in some months to between 5 and 6 billion per month at the peak in the summer and fall of 2012. To understand the evolution of costs, we consider some counterfactual experiments. As we showed in Section 6.1, two variables, application volume and a time trend, account for most of the variation in the price of intermediation,  $\phi$ , and so one important question is how much of the variation in intermediation expenditures is explained by these two factors. We use our regression estimates from column (6) of Table 4 to calculate predicted values of  $\phi$  under three sets of assumptions: (1)  $\phi$  does not react to applications; (2) the time trend was zero; and (3) the combination of (1) and (2). In each case, we assume that the level of  $\phi$  at the beginning of our sample period was at its actual level. We then use those estimates to calculate counterfactual expenditures on intermediation displayed in the figure, holding the volume of new loan originations fixed at its actual level.

The line labeled “No effect of apps” shows that if  $\phi$  had not responded to application volume, this would have lowered borrowers’ spending on intermediation by almost 50 billion dollars, from 147 billion to 101 billion. If, on the other hand, we let  $\phi$  vary with applications as it did but there was no increasing time trend, borrowers’ costs would have fallen by more than 90 billion dollars. Finally, the combination of the two counterfactual assumptions wipes out most of the increased cost, reducing it to only 12 billion over the 73 months of our sample. Thus, the upward trend in the price of intermediation documented in Section 6.1 resulted in a nearly three-fold increase in intermediation expenditures during this period. Capacity constraints also led to substantial increases in expenditures on intermediation.

Given our decomposition of  $\phi$  into origination costs ( $\pi$ ) and the PV of servicing costs (see Section 6.1), we can also study how much of the expenses were due to variation in the two components. In this case, our counterfactual experiment is to assume that servicing

---

<sup>34</sup>This therefore assumes that other types of mortgages, such as FHA loans or loans held on bank balance sheets, incur the same price of intermediation.

costs remained constant and that  $\pi$  evolved as predicted by the estimates in column (2) of the top panel of Table 6. Under these assumptions, origination costs led to \$73 billion in expenditures over the sample period or roughly half of all expenditure on intermediation, with increases in the net cost of servicing accounting for the remainder.

A natural question is whether it is reasonable to use  $\phi$  in October 2008 as the “reference level” in our counterfactuals given that it is substantially lower than what we observe over the rest of our sample period. While we do not have the Optimal Blue data going back further, a similar calculation based on the Freddie Mac survey rate implies that  $\phi$  was roughly constant at this low level over 2006–2008 (see Appendix D). Thus, it is not the case that the starting point of our  $\phi$  series is “unusual” relative to historical values.

Importantly, we are not making any welfare statements. What the numbers above reflect is a transfer of the surplus generated by lower rates (higher MBS prices) from households to intermediaries, taking the volume of loan originations as fixed. This is not to say that the high price of intermediation could not have welfare effects. In particular, had  $\phi$  remained at its lower original level, this would have led to additional mortgage originations that would have generated surplus for both borrowers and intermediaries.<sup>35</sup> Thus, if for instance post-crisis regulations led to the positive time trend in  $\phi$ , then one can argue that there was in fact a deadweight loss. That said, these regulations may of course have offsetting benefits, such as fewer mortgage defaults, that would have to be weighed in a full welfare analysis.

## 7.2 Effect on Mortgage Rates

In this section, we clarify how our measure relates to alternatives based on mortgage rates, and especially to the spread between rates and yields on MBS. Consider an intermediary similar to that in Section 2.1 that makes a loan with note rate  $r^n$  and pays the borrower  $p_{YSP}^n$ . However, rather than selling the loan, the intermediary finances  $p_{YSP}^n$  in the secondary market by issuing debt. Denoting the cost of funds on the debt by  $y^n$ , then the following equation relates the rate-spread,  $r^n - y^n$ , and the price of intermediation:

$$\phi^n = (r^n - y^n) \times \text{mult}_{MBS}^n. \quad (7)$$

Appendix F provides more detail on the derivation of equation (7) but the intuition is straightforward. The intermediary can choose between a single upfront payment  $\phi^n$  or a flow of income  $r^n - y^n$ . The present value of the flow of income, determined by the valuation multiple  $\text{mult}_{MBS}^n$  (see Section 2.3.1) must, by arbitrage, equal the single upfront payment.

---

<sup>35</sup>A full discussion of the strength of this channel requires a model of borrower behavior and is beyond the scope of this paper.

If we solve equation (7) for the note rate,

$$r^n = y^n + \phi^n / \text{mult}_{MBS}^n, \quad (8)$$

we can use our counterfactual  $\phi$  series from the previous section to generate a counterfactual interest rate series as displayed in the bottom panel of Figure 9. The figure shows that the costs of intermediation did have a significant effect on the interest rates. Our experiment of setting the sensitivity of the price of intermediation to applications to zero would have reduced rates by 12 basis points on average and as much as 95 basis points. Eliminating the time trend would have reduced Rate101 by 17 basis points on average and by 28 basis points by the end of the sample. The combination of the two counterfactual experiments implies a decrease in Rate101 of 30–40 basis points over an extended period in 2012–2013.

Equation (8) also allows us to relate our measure of the price of intermediation to other treatments of the topic. Specifically,  $r^n - y^n$  is similar to what market participants normally refer to as the “primary-secondary spread” which is often viewed as a measure of the price of intermediation.<sup>36</sup> Equation (8) illustrates the challenge of using  $r^n - y^n$  to measure changes in the price of intermediation. Changes in the value of the interest strip can change rates even when there is no change in  $\phi$ . Intuitively, suppose that market expectations of prepayments go down, leading the MBS multiplier to go up. Even if the cost of funds  $y^n$  and the price of intermediation  $\phi^n$  stay the same, the primary market rate  $r^n$  will fall, leading one to conclude, incorrectly, that the price of intermediation has fallen. The reason for the fall in rates is that the price of intermediation is now spread out over a longer period.

Figure 10 illustrates the relationship between the rate spread,  $r - y$ , and the price of intermediation,  $\phi$ . At some points in our sample, the rate spread and  $\phi$  tell similar stories. For example, in 2012, both  $\phi$  and the rate spread almost doubled, a fact that is not surprising given that the multiplier, shown in the bottom part of the panel, is relatively stable. But at other points, the rate spread and  $\phi$  give different accounts. For example, in late 2008 and early 2009, the rate spread spikes and then drops dramatically. The decline does not show up in  $\phi$ , which remains within a relatively narrow band. The difference, of course, is the multiplier, which rises from one at the end of 2008 to nearly six at the beginning of 2010.

Note that the strong time trend we documented in Section 6.1 does not show up in the rate spread. The difference between the trends in the two series is explained by the rising

---

<sup>36</sup>Analysts typically define the primary-secondary spread as the difference between an index of primary market rates like the 30-year fixed rate index from the Freddie Mac Primary Mortgage Market Survey and compare it with the yield on a MBS trading at par. This would only be equivalent to  $r^n - y^n$  if the rebate on the Freddie Mac loans was zero, which is typically not the case. Since the rebates are usually positive, the primary-secondary spread overstates the rate spread. Even if intermediation was free, a positive rebate would imply a positive spread between the note rate and yield on a loan trading at par.



multiplier, which ensured that the increase in the price of intermediation translated into a smaller increase in rates. Put differently, if the multiplier had remained constant over this period, the rate spread would have widened much more dramatically than it did.

## 8 Conclusion

Over the 2000–2014 period, residential mortgage originations in the United States averaged about 2.2 trillion dollars per year ([Inside Mortgage Finance, 2016](#)). Given the size of the market, even relatively small changes in the price of intermediation, defined as the difference between the value of a mortgage’s cash flows to investors in the MBS market and what a borrower receives in funding, add up to significant changes in implicit borrower costs. In this paper, we have systematically studied how the price of intermediation has evolved over the 2008–2014 period, during which monetary policy and other macroeconomic and financial factors led to record low interest rates.

We find a significant upward trend in the price of intermediation that appears to be driven by increased net costs of mortgage servicing and potentially the increased labor intensity of mortgage underwriting due to new regulations and lenders being more averse to liability risk. In addition, there are substantial fluctuations around this trend that are closely related to the level of demand for new mortgages; this finding is indicative of capacity constraints in mortgage originations. We estimate that the upward trend and the sensitivity to demand combined to raise intermediation costs for borrowers by about 135 billion dollars over the 73 months in our sample, holding origination volumes constant.

About two-thirds of household expenditures on intermediation over this period, or about 95 billion dollars, involved refinances of existing mortgages. One can interpret this figure as one of the costs of having a mortgage market that is dominated by fixed-rate mortgages, since mortgagors need to refinance to benefit from lower rates. Expenditures on intermediation could potentially be much lower if refinancing did not require a full re-underwriting of the mortgage even though the reduction in the monthly payment will generally lead to a reduced credit risk relative to the outstanding loan.<sup>37</sup> Streamlining the process would not only save borrowers time, but potentially also substantial amounts of money. Another way to reduce borrowers’ intermediation expenses would be to have more adjustable-rate mortgages, since the transmission of interest rate changes to household cash flow in that case requires no action by financial intermediaries.

Capacity constraints in mortgage intermediation have affected the transmission of monetary policy. Even though we show that offers to borrowers in the primary mortgage market

---

<sup>37</sup>Credit risk may not be lower if the borrower withdraws equity in a cash-out refinance.

generally respond very quickly to MBS price changes, we also find that the pass-through of MBS price increases is subdued when application volumes are high. This result implies that policy actions that increase MBS prices—for instance, an announcement of increased QE asset purchases—will pass through less strongly if demand for mortgages was already high. Put differently, in such a situation, intermediaries may absorb a large chunk of the increase in MBS values, rather than passing it on to borrowers. Thus, intermediation frictions are important for policymakers to consider when designing policy actions that primarily target the mortgage market.

## References

- Allen, Jason, Robert Clark, and Jean-François Houde. 2014. “Price Dispersion in Mortgage Markets.” *Journal of Industrial Economics* 62(3): 377–416.
- Ausubel, Lawrence M. 1991. “The Failure of Competition in the Credit Card Market.” *American Economic Review* 81(1): 50–81.
- Bachmeier, Lance J, and James M Griffin. 2003. “New Evidence on Asymmetric Gasoline Price Responses.” *Review of Economics and Statistics* 85(3): 772–776.
- Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. 2015. “Regional Heterogeneity and Monetary Policy.” Staff Report 731. Federal Reserve Bank of New York.
- Bernanke, Ben S., and Mark Gertler. 1995. “Inside the Black Box: The Credit Channel of Monetary Policy Transmission.” *Journal of Economic Perspectives* 9(4): 27–48.
- Bernanke, Ben S., Mark Gertler, and Simon Gilchrist. 1999. “The Financial Accelerator in a Quantitative Business Cycle Framework.” In *Handbook of Macroeconomics*, eds. John B. Taylor and Michael Woodford, vol. 1, 1341–1393. Amsterdam: Elsevier.
- Borenstein, Severin, A. Colin Cameron, and Richard Gilbert. 1997. “Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes?” *Quarterly Journal of Economics* 112(1): 305–339.
- Boyarchenko, Nina, Andreas Fuster, and David O. Lucca. 2015. “Understanding Mortgage Spreads.” Staff Report 674. Federal Reserve Bank of New York.
- Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2017. “Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks.” Working Paper 23288. National Bureau of Economic Research.
- Calem, Paul S., Michael B. Gordy, and Loretta J. Mester. 2006. “Switching Costs and Adverse Selection in the Market for Credit Cards: New evidence.” *Journal of Banking & Finance* 30(6): 1653–1685.
- Cordell, Larry, Karen Dynan, Andreas Lehnert, Nellie Liang, and Eileen Mauskopf. 2009. “Designing Loan Modifications to Address the Mortgage Crisis and the Making Home Affordable Program.” Finance and Economics Discussion Series 2009-43. Washington, DC: Federal Reserve Board.
- Di Maggio, Marco, Amir Kermani, and Christopher Palmer. 2016. “Unconventional Monetary Policy and the Allocation of Credit.” Working Paper. UC Berkeley.
- Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2015. “The Deposits Channel of Monetary Policy.” Working Paper. New York: New York University.
- Driscoll, John C., and Ruth A. Judson. 2013. “Sticky Deposit Rates.” Finance and Economics Discussion Series 2013-80. Washington, DC: Federal Reserve Board.

- Fratantoni, Michael. 2016. "Mortgage Lender Profitability and Its Impact on Market Spreads." *Journal of Structured Finance* 22(1): 69–76.
- Fuster, Andreas, Laurie Goodman, David Lucca, Laurel Madar, Linsey Molloy, and Paul Willen. 2013. "The Rising Gap between Primary and Secondary Mortgage Rates." *Federal Reserve Bank of New York Economic Policy Review* 19(2): 17–39.
- Fuster, Andreas, and Paul S. Willen. 2010. "\$1.25 Trillion is Still Real Money: Some Facts About the Effects of the Federal Reserve's Mortgage Market Investments." Public Policy Discussion Paper No. 10-4. Boston: Federal Reserve Bank of Boston.
- Gabaix, Xavier, David Laibson, Deyuan Li, Hongyi Li, Sidney Resnick, and Casper G de Vries. 2016. "The Impact of Competition on Prices with Numerous Firms." *Journal of Economic Theory* 165: 1–24.
- Gertler, Mark, and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics* 7(1): 44–76.
- Gilchrist, Simon, David López-Salido, and Egon Zakrajšek. 2015. "Monetary Policy and Real Borrowing Costs at the Zero Lower Bound." *American Economic Journal: Macroeconomics* 7(1): 77–109.
- Gilchrist, Simon, and Egon Zakrajšek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review* 102(4): 1692–1720.
- Goodman, Laurie S., and Jun Zhu. 2013. "Reps and Warrants: Lessons from the GSEs Experience." Housing Finance Policy Center Research Report. Urban Institute.
- Greenwood, Robin, and David Scharfstein. 2013. "The Growth of Finance." *The Journal of Economic Perspectives* 27(2): 3–28.
- Hancock, Diana, and Wayne Passmore. 2011. "Did the Federal Reserve's MBS Purchase Program Lower Mortgage Rates?" *Journal of Monetary Economics* 58: 498–514.
- Hancock, Diana, and Wayne Passmore. 2015. "How Does the Federal Reserve's Large-Scale Asset Purchases (LSAPs) Influence Mortgage-Backed Securities (MBS) Yields and U.S. Mortgage Rates?" *Real Estate Economics* 43(4): 855–890.
- Hannan, Timothy H., and Allen N. Berger. 1991. "The Rigidity of Prices: Evidence from the Banking Industry." *American Economic Review* 81(4): 938–945.
- Hendricks, Bradley E., Jed J. Neilson, Catherine Shakespeare, and Christopher D. Williams. 2016. "Responding to Regulatory Uncertainty: Evidence from Basel III." Working Paper No. 1213. Ann Arbor: University of Michigan Ross School of Business.
- Hurst, Erik, Benjamin Keys, Amit Seru, and Joseph Vavra. 2016. "Regional Redistribution through the US Mortgage Market." *American Economic Review* 106(10): 2982–3028.
- Inside Mortgage Finance. 2016. *Mortgage Market Statistical Annual 2016*.

- Kashyap, Anil K, and Jeremy C. Stein. 2000. "What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?" *American Economic Review* 90(3): 407–428.
- Kaul, Karan, and Laurie Goodman. 2016. "Nonbank Servicer Regulation: New Capital and Liquidity Requirements Don't Offer Enough Loss Protection." Housing Finance Policy Center Brief. Washington, DC: Urban Institute.
- Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2011. "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity* (2): 215–287.
- Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2013. "The Ins and Outs of LSAPs." Proceedings of the Federal Reserve Bank of Kansas City's Jackson Hole Symposium on the Global Dimensions of Unconventional Monetary Policy.
- Malamud, Semyon, and Andreas Schrimpf. 2017. "Intermediation Markups and Monetary Policy Passthrough." Research Paper 16-75. Swiss Finance Institute.
- Maturana, Gonzalo. 2015. "When Are Modifications of Securitized Loans Beneficial to Investors?" Working Paper. Atlanta: Emory University.
- Mora, Nada. 2014. "The Weakened Transmission of Monetary Policy to Consumer Loan Rates." *Federal Reserve Bank of Kansas City Economic Review* 99(1): 93–117.
- Neumark, David, and Steven A. Sharpe. 1992. "Market Structure and the Nature of Price Rigidity: Evidence from the Market for Consumer Deposits." *Quarterly Journal of Economics* 107(2): 657–680.
- Peltzman, Sam. 2000. "Prices Rise Faster than They Fall." *Journal of Political Economy* 108(3): 466–502.
- Scharfstein, David S., and Adi Sunderam. 2013. "Concentration in Mortgage Lending, Refinancing Activity and Mortgage Rates." Working Paper 19156. National Bureau of Economic Research.
- Sharpe, Steve A., and Shane M. Sherlund. 2016. "Crowding Out Effects of Refinancing on New Purchase Mortgages." *Review of Industrial Organization* 48(2): 209–239.
- Stroebel, Johannes, and John B. Taylor. 2012. "Estimated Impact of the Federal Reserve's Mortgage-Backed Securities Purchase Program." *International Journal of Central Banking* 8(2): 1–42.
- Vickery, James, and Joshua Wright. 2013. "TBA Trading and Liquidity in the Agency MBS Market." *Federal Reserve Bank of New York Economic Policy Review* 19(2): 1–18.
- Walentin, Karl. 2014. "Business Cycle Implications of Mortgage Spreads." *Journal of Monetary Economics* 67: 62–77.
- Yankov, Vladimir. 2014. "In Search of a Risk-free Asset." Finance and Economics Discussion Series 2014-108. Washington, DC: Federal Reserve Board.

<b>Loan-level Characteristics for Baseline Scenario</b>			
<b>Variable</b>	<b>Value</b>	<b>Variable</b>	<b>Value</b>
Loan amount	300,000	Prepayment penalty	None
Loan-to-value ratio (LTV)	80	Lock period	30 days
FICO	750	Owner-occupied or investment	Owner-occupied
Program	Conforming	State	CA
Loan type	Fixed	MSA	Los Angeles
Term	30 year		

<b>Rate Sheet Information</b>						
	<b>Mean</b>	<b>SD</b>	<b>p10</b>	<b>p50</b>	<b>p90</b>	
Unique lenders	63					
Lenders per day	21.7	5.6	13	21	28	
Offers per ratesheet	22.9	18.5	9	16	48	
Rebate offered	1.2	2.2	-1.8	1.5	3.7	
Rate offered	4.55	0.71	3.63	4.50	5.50	

Table 1: Characteristics of the Proprietary Lender Ratesheet Data from Optimal Blue Used for Our Analysis

Notes: The rate/point offers depend on borrower- and loan-level characteristics, such as the FICO score and loan type. The baseline scenario used in our analysis is given in the top panel. The bottom panel shows descriptive statistics on the number of lenders and the offers per rate sheet in our data, as well as on the distributions of rebates and rates in the pooled sample.

	(1)	(2)	(3)	(4)	(5)
$\Delta p_{TBA,t}$	0.915*** (0.014)				
$\Delta p_{TBA,t}^+$		0.795*** (0.023)	0.790*** (0.024)	0.781*** (0.022)	0.776*** (0.022)
$\Delta p_{TBA,t}^-$		1.021*** (0.022)	1.027*** (0.022)	1.024*** (0.022)	1.029*** (0.022)
$\Delta p_{TBA,t-1}^+$			0.078*** (0.020)		0.088*** (0.019)
$\Delta p_{TBA,t-1}^-$			0.012 (0.023)		0.004 (0.023)
$\Delta p_{TBA}^+ \times \text{Applications}_{t-1}$				-0.096*** (0.023)	-0.106*** (0.024)
$\Delta p_{TBA}^- \times \text{Applications}_{t-1}$				0.010 (0.022)	0.021 (0.022)
$\Delta p_{TBA,t-1}^+ \times \text{Applications}_{t-1}$					-0.021 (0.018)
$\Delta p_{TBA,t-1}^- \times \text{Applications}_{t-1}$					-0.034 (0.023)
$\text{Applications}_{t-1}$				0.002 (0.004)	0.003 (0.005)
Constant	-0.001 (0.002)	0.022*** (0.004)	0.016*** (0.005)	0.023*** (0.004)	0.015*** (0.005)
Adj. R <sup>2</sup>	0.88	0.89	0.89	0.89	0.89
Observations	1431	1431	1431	1431	1431

Robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 3: Regressions of  $\Delta p_{YSP}^{Rate101}$  on  $\Delta p_{TBA}^{Rate101}$  at Daily Frequency

Notes: “ $\text{Applications}_{t-1}$ ” is the daily HMDA application volume from the previous business day, normalized to have a mean of zero and a standard deviation of one.  $\Delta p_{TBA}^+$  indicates the magnitude of the day-over-day change in TBA price for positive changes only, and  $\Delta p_{TBA}^-$  corresponds to the negative changes. Columns 3 and 5 also include a price change lagged by one business day. All prices are normalized to the same rate (“Rate101”).

	$\phi$ , OLS				$\phi$ , IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Applications	0.199* (0.108)	0.357*** (0.042)	0.331*** (0.049)	0.275*** (0.062)	0.204 (0.137)	0.379*** (0.053)	0.407*** (0.058)	0.308*** (0.094)
Time Trend (monthly)		0.026*** (0.002)	0.016** (0.008)	0.013 (0.008)		0.026*** (0.002)	0.025*** (0.007)	0.016 (0.011)
Volatility			-0.094 (0.077)	-0.023 (0.075)			-0.027 (0.074)	-0.016 (0.070)
Lender Conc.			0.060 (0.085)	0.060 (0.078)			0.148** (0.069)	0.086 (0.090)
R.E. Payroll			0.199*** (0.063)	0.207*** (0.051)			0.154** (0.063)	0.192*** (0.066)
Rate101				-0.196 (0.119)				-0.160 (0.151)
Constant	1.398*** (0.124)	-14.485*** (1.319)	-8.725* (4.981)	-5.827 (5.231)	1.398*** (0.124)	-14.677*** (1.331)	-14.263*** (4.578)	-7.999 (7.049)
Obs.	73	73	73	73	73	73	73	73
Adj. R <sup>2</sup>	0.10	0.84	0.87	0.88				
F-stat					95.5	93.8	37.6	29.3

Newey-West standard errors (4 lags) in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### Predicted vs. Actual, Specification (2)

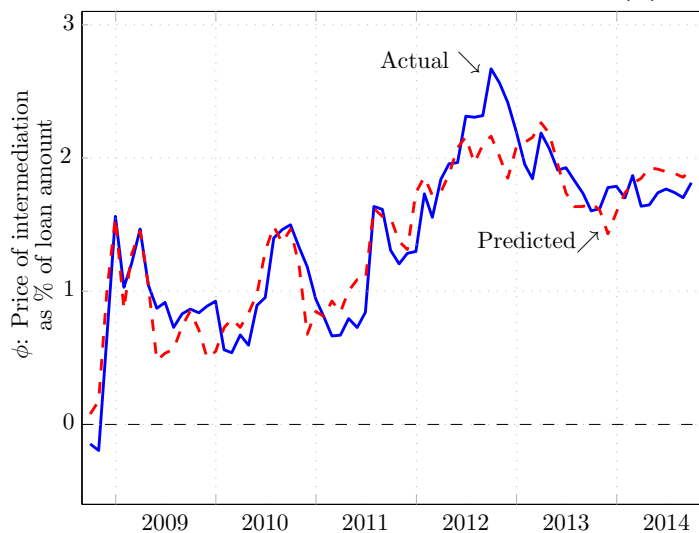


Table 4: Understanding Time-Variation in the Price of Intermediation: Regression in Levels  
Notes: The definition of  $\phi$  and the explanatory variables can be found in Section 3. All explanatory variables except for the time trend (where the unit is calendar month) have been standardized over the relevant sample, so that the coefficients can be interpreted as the effect of one standard deviation of the relevant variable on the price of intermediation.  $R^2$  is not provided for instrumental variables (IV) specification since it does not have a natural interpretation in this case.



	$\Delta\phi$ , OLS			$\Delta\phi$ , IV		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta$ Applications	0.247*** (0.046)	0.234*** (0.056)	0.259*** (0.064)	0.312*** (0.048)	0.313*** (0.054)	0.491*** (0.070)
$\Delta$ Volatility		-0.005 (0.104)	-0.023 (0.109)		0.060 (0.109)	0.079 (0.080)
$\Delta$ Concentration		-0.081 (0.078)	-0.096 (0.077)		-0.037 (0.079)	-0.034 (0.091)
$\Delta$ R.E. Payroll		0.032 (0.053)	0.034 (0.055)		0.028 (0.057)	0.028 (0.072)
$\Delta$ Rate101			0.179 (0.195)			0.582*** (0.197)
Constant	0.027 (0.023)	0.024 (0.024)	0.029 (0.026)	0.027 (0.023)	0.029 (0.024)	0.050* (0.028)
Obs.	72	72	72	72	72	72
Adj. R <sup>2</sup>	0.40	0.39	0.39			
First-stage F-stat				85.8	65.9	28.6

Robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 5: Understanding Time-Variation in the Price of Intermediation: Regression in Monthly Changes

Notes: The definitions of  $\phi$ ,  $\pi$  and the explanatory variables can be found in Section 3. As in Table 4, all right-hand side variables are standardized (before taking monthly changes).

	$\pi$			$\phi - \pi$		
	(1)	(2)	(3)	(4)	(5)	(6)
Applications	0.336*** (0.089)	0.407*** (0.068)	0.388*** (0.068)	-0.132 (0.086)	-0.027 (0.048)	0.020 (0.046)
Time Trend (monthly)		0.010*** (0.003)	0.006 (0.009)		0.015*** (0.003)	0.019** (0.008)
Volatility			0.161** (0.073)			-0.187*** (0.055)
Concentration			0.033 (0.101)			0.116 (0.080)
R.E. Payroll			0.296*** (0.068)			-0.142** (0.058)
Constant	1.132*** (0.080)	-5.341*** (1.896)	-2.522 (5.347)	0.265*** (0.085)	-9.336*** (1.744)	-11.741** (4.785)
Obs.	73	73	73	73	73	73
First-stage F-stat	95.5	93.8	37.6	95.5	93.8	37.6

Newey-West standard errors (4 lags) in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

	$\Delta \pi$		$\Delta (\phi - \pi)$	
	(1)	(2)	(3)	(4)
$\Delta$ Applications	0.389*** (0.069)	0.430*** (0.068)	-0.077** (0.038)	-0.117*** (0.031)
$\Delta$ Volatility		0.234* (0.121)		-0.174*** (0.047)
$\Delta$ Concentration		-0.019 (0.089)		-0.018 (0.047)
$\Delta$ R.E. Payroll		0.059 (0.073)		-0.031 (0.029)
Constant	0.011 (0.027)	0.022 (0.027)	0.016 (0.013)	0.007 (0.013)
Obs.	72	72	72	72
First-stage F-stat	85.8	65.9	85.8	65.9

Robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 6: Decomposing Time-Variation in the Price of Intermediation

Notes:  $\pi$  measures origination costs while  $\phi - \pi$  measures the present value of the net cost of servicing, as discussed in Section 2.3.1. All regressions use the gap between weighted average coupon on outstanding conventional MBS and the ten-year Treasury yield as instrumental variable for applications, as discussed in Section 6.1.

	(1)	(2)	(3)	(4)
	Delay	Delay	$\Delta$ Delay	$\Delta$ Delay
Std. New Applications	3.672*** (0.681)	4.034*** (0.643)		
Time Trend	0.148*** (0.0404)	0.150*** (0.0376)		
$\Delta$ Std. New Applications			2.235*** (0.485)	2.830*** (0.328)
Dec. Indicator				1.564** (0.653)
Jan. Indicator				-4.126*** (0.905)
Constant	-54.28** (25.29)	-57.25** (23.93)	0.165 (0.212)	0.381** (0.161)
Month FEs	no	yes	no	no
Adj. R <sup>2</sup>	0.53	0.55	0.38	0.61
Observations	81	81	80	80

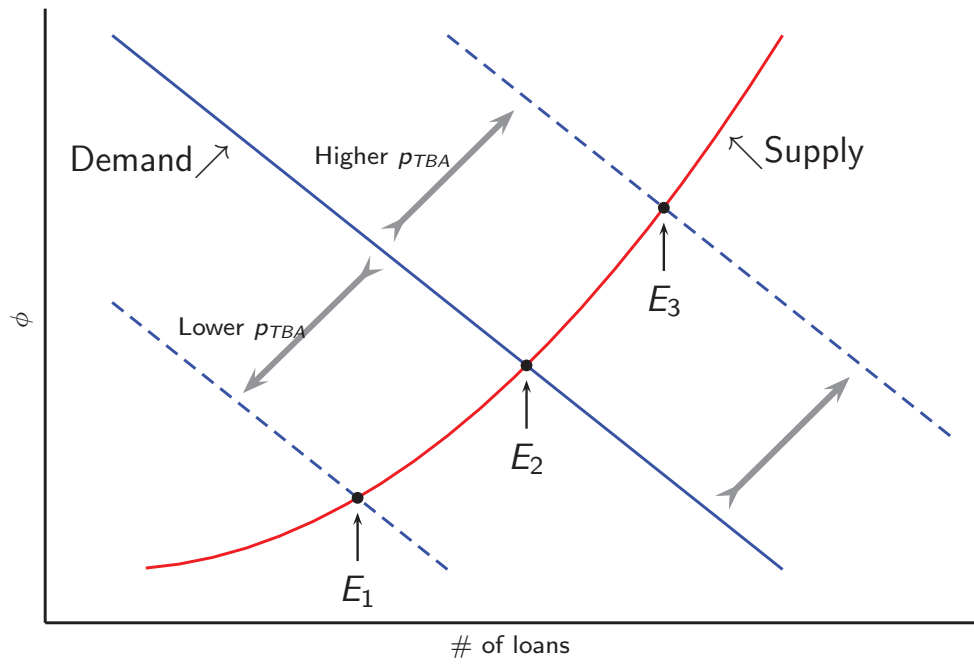
Robust standard errors shown parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 7: Regression of HMDA Processing Delays on Volume of Loan Applications

Notes: Monthly regression using data from January 2008 to September 2014. The dependent variable, delay, is defined as the number of days between loan application and action (accept, withdraw, or deny); we end in September 2014 rather than October (as for the rest of the analysis) since the October delay is downward biased due to censoring (only applications processed by the end of 2014 are in our data). We measure loan demand using the count of loan applications from HMDA divided by the number of business days in a month; we then normalize the variable to have a mean of zero and a standard deviation of one. Newey-West robust standard errors (four lags) shown for Columns (1) and (2); robust standard errors shown for Columns (3) and (4). Column (2) contains calendar-month fixed effects to control for potential seasonality.

### A. Supply and Demand for Intermediation



### B. Perfectly Elastic Supply

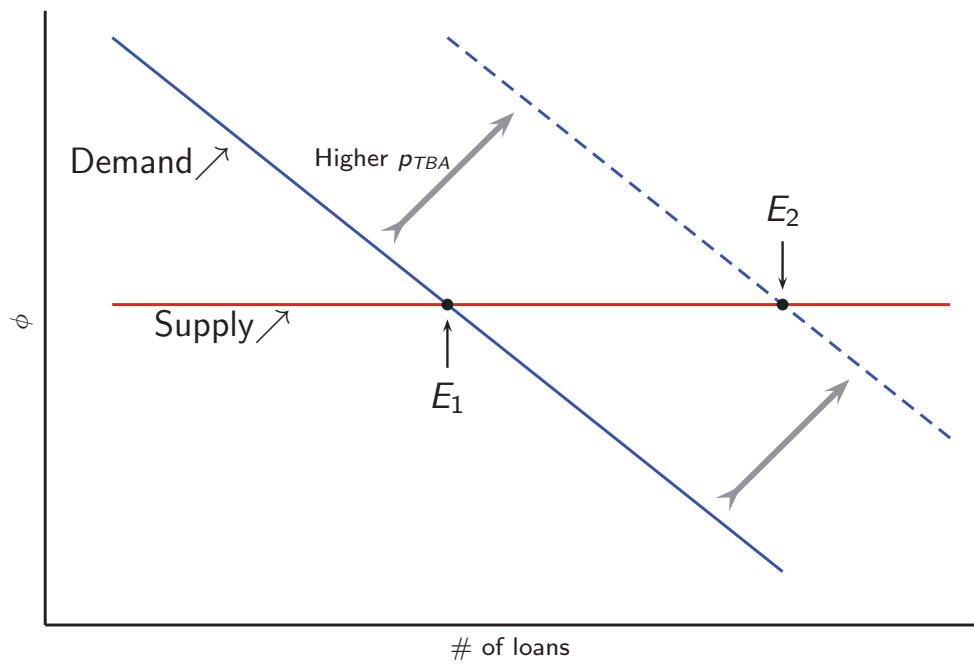


Figure 1: The Market for Financial Intermediation in the Mortgage Market

Visit our website at [www.53.com/wholesalemortgage](http://www.53.com/wholesalemortgage)  
 E-mail: WholesaleRegistration.Cincinnati@53.com  
 To Register and Lock Rates Online:  
 Go to [www.53.com/wholesalemortgage](http://www.53.com/wholesalemortgage)  
 and select "Wholesale Connect".  
 Contact Your 5/3 Account Executive to gain access.

Group 7 State(s): IL



Wholesale Rate Sheet For Broker Use Only  
 Traditional and Table Fund  
 Rate Sheet Number: 2010-1618  
 Effective @8:30 AM EST of: 6/7/2010

Maximum Combined Broker Compensation (YSP and Broker Fees)  
 Loan Amount <= \$417,000 = 3.5% of Total Loan Amount  
 Loan Amounts > \$417,000 = 2.5% of Total Loan Amount

Off Sheet Pricing available: Contact your Account Executive for information  
 FREDDIE MAC OPEN ACCESS NOW AVAILABLE - SEE PAGE 5 FOR PRICING ADJUSTMENTS

Conforming Fixed Products - Upto FHLMC Limit

21-30 Yr. Fixed (FX30)			
	15 day	30 day	60 day
5.375	(4.700)	(4.575)	(4.450)
5.250	(3.950)	(3.825)	(3.700)
5.125	(3.325)	(3.200)	(2.950)
5.000	(2.825)	(2.700)	(2.450)
4.875	(2.325)	(2.200)	(1.950)
4.750	(1.575)	(1.450)	(1.200)
4.625	(0.825)	(0.700)	(0.450)
4.600	(0.325)	(0.200)	0.050
4.375	0.300	0.425	0.675
4.250	1.175	1.300	1.550
4.125	2.675	2.800	3.175
4.000	3.425	3.550	3.800
3.875	4.050	4.175	4.550
3.750	4.925	5.050	5.300

16-20 Yr. Fixed (FX20)			
	15 day	30 day	60 day
5.375	(4.820)	(4.695)	(4.445)
5.250	(4.195)	(4.070)	(3.820)
5.125	(3.820)	(3.695)	(3.445)
5.000	(3.445)	(3.320)	(3.070)
4.875	(2.945)	(2.820)	(2.695)
4.750	(2.320)	(2.195)	(2.070)
4.625	(1.820)	(1.695)	(1.570)
4.600	(1.445)	(1.320)	(1.070)
4.375	(0.945)	(0.820)	(0.570)
4.250	(0.320)	(0.195)	0.055
4.125	3.680	3.605	4.055
4.000	4.180	4.305	4.660
3.875	4.805	4.930	5.180
3.750	5.430	5.555	5.805

10-15 Yr. Fixed (FX15)			
	15 day	30 day	60 day
5.125	(4.520)	(4.395)	(4.270)
5.000	(4.270)	(4.145)	(3.895)
4.875	(3.895)	(3.770)	(3.645)
4.750	(3.395)	(3.270)	(3.145)
4.625	(2.895)	(2.770)	(2.520)
4.600	(2.520)	(2.395)	(2.270)
4.375	(2.145)	(2.020)	(1.895)
4.250	(1.845)	(1.520)	(1.395)
4.125	(0.770)	(0.645)	(0.395)
4.000	(0.395)	(0.270)	(0.020)
3.875	0.105	0.230	0.480
3.750	0.730	0.855	0.980
3.625	1.855	1.980	2.230
3.500	2.480	2.805	2.750

AGENCY JUMBO LOANS

PRICE ADJUSTMENTS - FOR APPROVED MSA'S ONLY				
Product	Purpose	LTV %		
		<=75	>75 & <=80	>80
Fixed PR	Purchase R/T Refi	0.750	0.750	0.750
	Cash Out Refi	1.750	N/A	N/A
*Fixed IO	Purchase R/T Refi	0.750	0.750	0.750
	Cash Out Refi	1.750	N/A	N/A
*ARM IO	Purchase R/T Refi	1.500	2.250	2.250
	Cash Out Refi	2.500	N/A	N/A

All pricing bumps per Conforming products such as Credit Score/LTV,  
 Cash Out, Escrow Waiver & others apply in addition to above adjustments  
 \*IO Only available for apps prior to 5/17/10. Must Close by 6/13/10

Non-Agency Jumbo Fixed Products - Max Loan Amount \$1MM

Jumbo 16-30 Yr. Fixed (FJ30)			
	15 day	30 day	60 day
5.125	(3.255)	(3.130)	(2.880)
5.000	(3.130)	(3.005)	(2.755)
5.075	(3.005)	(2.880)	(2.630)
5.750	(2.505)	(2.380)	(2.130)
5.625	(2.005)	(1.880)	(1.630)
5.600	(1.505)	(1.380)	(1.130)
5.375	(1.130)	(1.005)	(0.755)
5.250	(0.505)	(0.380)	(0.130)
5.125	0.120	0.245	0.495
5.000	0.745	0.870	1.120
4.875	1.370	1.495	1.745
4.750	2.120	2.245	2.495

Jumbo 10-15 Yr. Fixed (FJ15)			
	15 day	30 day	60 day
5.625	(2.020)	(1.895)	(1.645)
5.500	(1.895)	(1.770)	(1.520)
5.375	(1.145)	(1.020)	(0.770)
5.250	(0.895)	(0.770)	(0.520)
5.125	(0.845)	(0.520)	(0.270)
5.000	(0.270)	(0.145)	(0.020)
4.875	0.355	0.480	0.605
4.750	0.730	0.855	0.980
4.625	1.105	1.230	1.480
4.500	1.480	1.605	1.855
4.375	2.230	2.355	2.480
4.250	2.730	2.855	2.980

ALL ADJUSTMENTS BELOW ARE TO PRICE EXCEPT IF NOTED OTHERWISE

Figure 2: An Example of a Rate Sheet

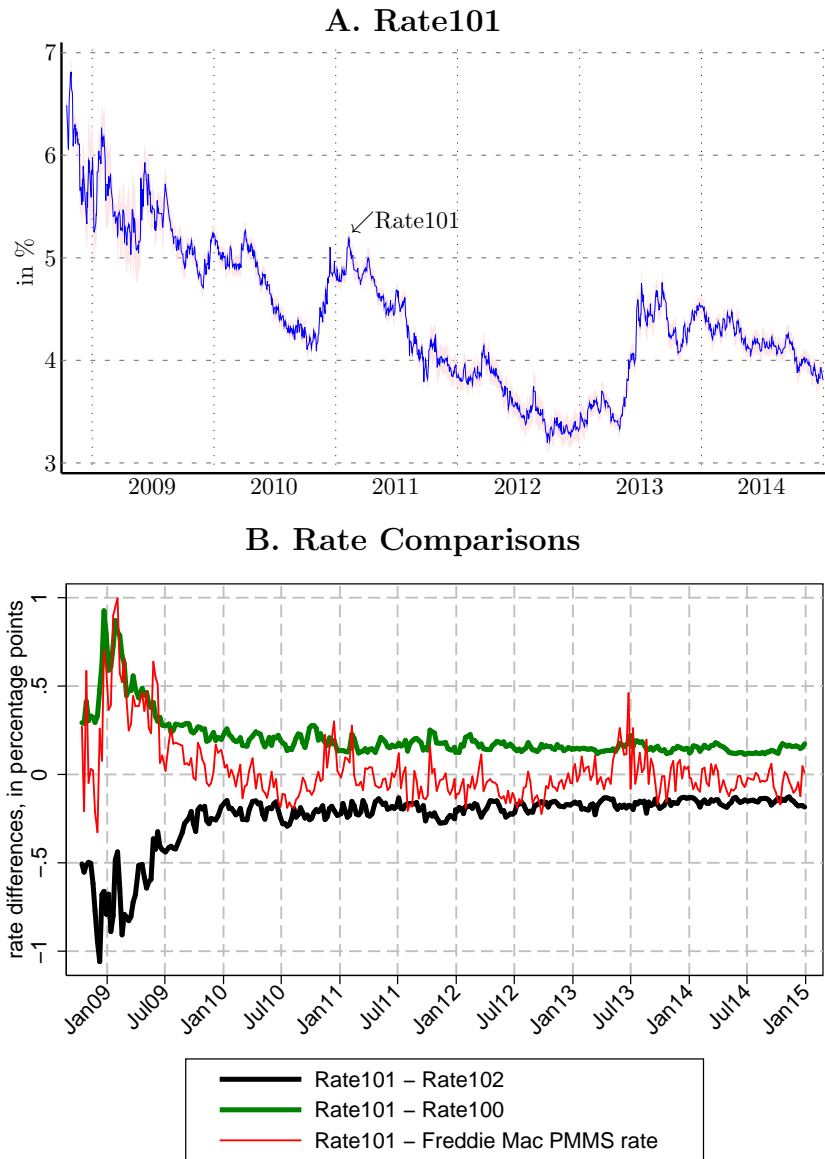


Figure 3: Mortgage Rates Over Our Sample Period

Notes: Top panel shows Rate101, the median rate across lenders at which a rebate of +1 point is offered (see Section 3.1.1). Bottom panel shows the difference between Rate101 and Rate102 (for a rebate of +2), Rate100 (for a rebate of zero), and the Freddie Mac Primary Mortgage Market Survey series for 30-year fixed-rate mortgages. All series are weekly averages. Shaded areas in panel A is standard deviation of all Rate101 offers in our data. For the Rate101-Freddie Mac series, the Rate101 series is restricted to the days that are reported in the weekly Freddie Mac series.

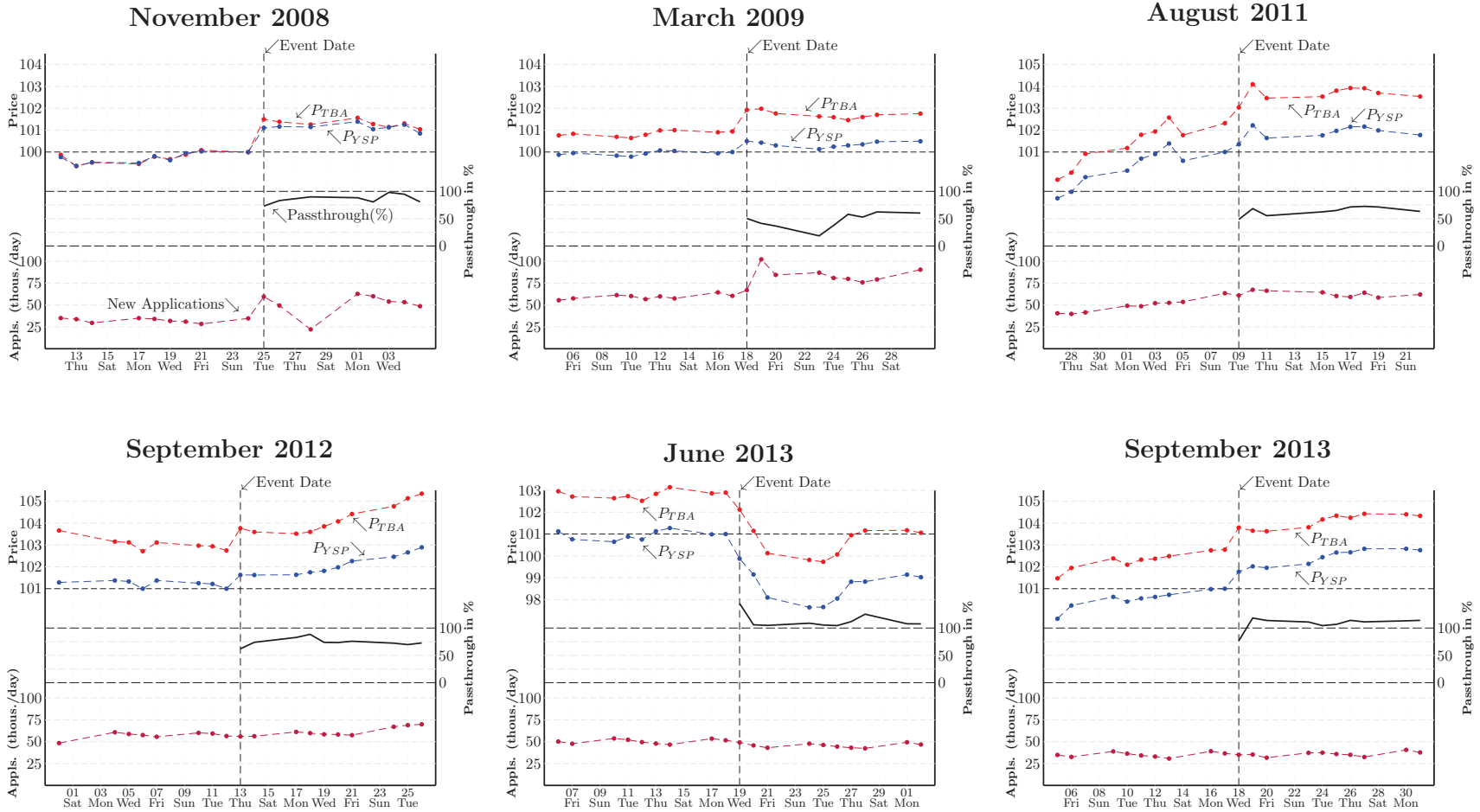
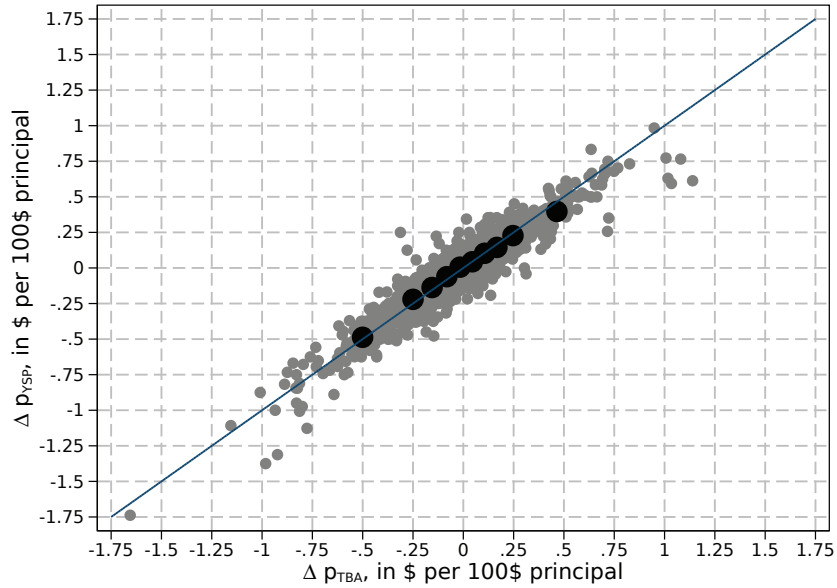


Figure 4: Event Studies of Major Monetary Policy (QE) Announcements Over Our Sample Period

Notes: Top lines show the TBA and YSP prices holding the rate fixed at Rate101 (Rate100 in 2008-2009) of the day prior to the announcement. Middle line shows the pass-through of the TBA price change to rebates (YSPs). Bottom line shows daily application counts as measured by HMDA (non-weekends/holidays only).

### A. Daily First Differences of $p_{YSP}$ and $p_{TBA}$



### B. Coefficients of First Difference Regression by Year

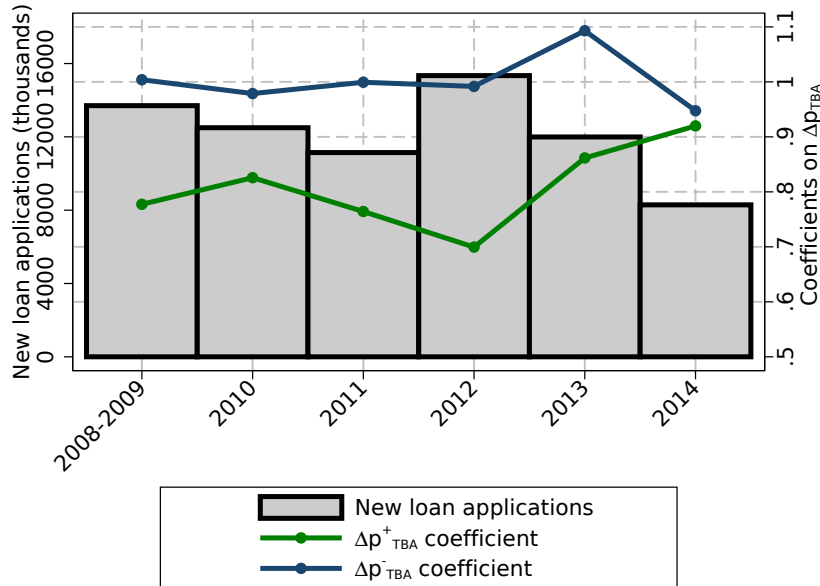
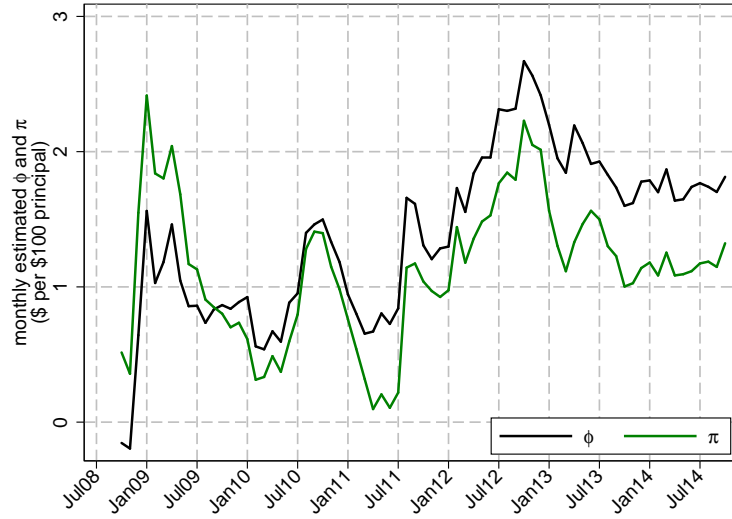


Figure 5: High Frequency Pass-Through

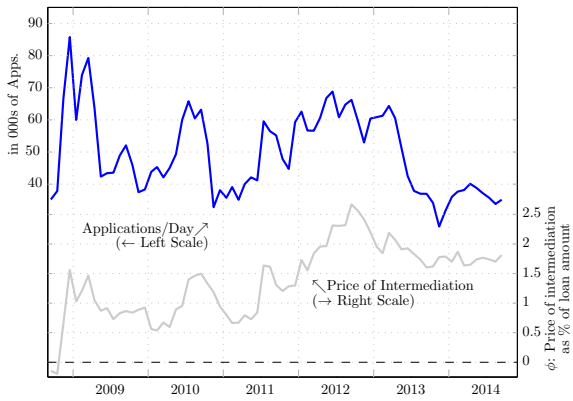
Notes: Top panel shows scatterplot of daily first differences of rebates (YSPs) and MBS prices. Grey dots show every data point (for the full sample period, October 2008 - October 2014) while black dots show averages by deciles of  $\Delta p_{TBA}$  (“binscatter”). Diagonal line is the 45-degree line, where points would lie with perfect pass-through. Both the YSP and MBS series are for Rate101, as discussed in Section 5. These data underlie the regressions in Table 3. Bottom panel shows coefficients of first difference regression ( $\Delta p_{YSP} = \beta_0 + \beta_+ \Delta p_{TBA}^+ + \beta_- \Delta p_{TBA}^- + \varepsilon$ ) run by year, plotted alongside bars that show the annual level of new mortgage applications in the HMDA data. Source: Optimal Blue; J.P. Morgan Markets; HMDA; authors’ calculations.



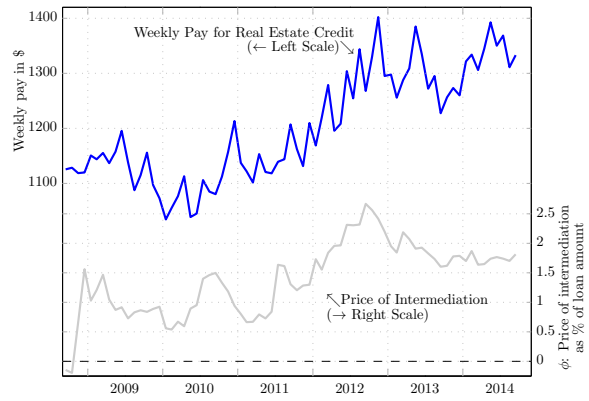
### A. Monthly Estimated $\phi$ and $\pi$



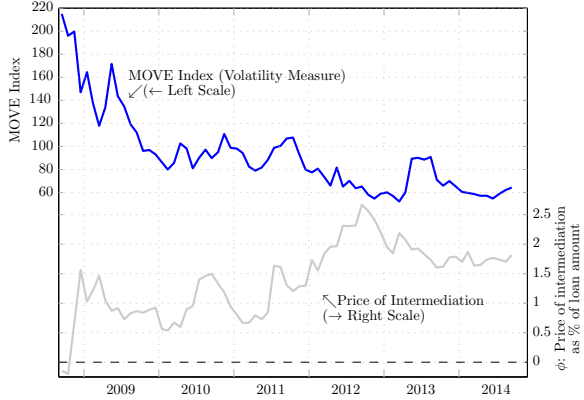
### B. $\phi$ vs. Loan Applications



### C. $\phi$ vs. R.E. Credit Wages



### D. $\phi$ vs. Interest Volatility



### E. $\phi$ vs. Concentration

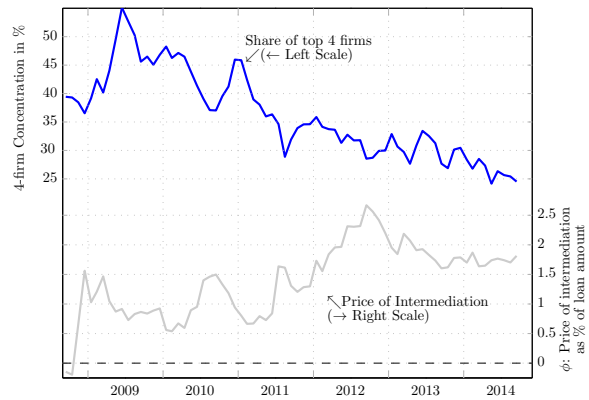
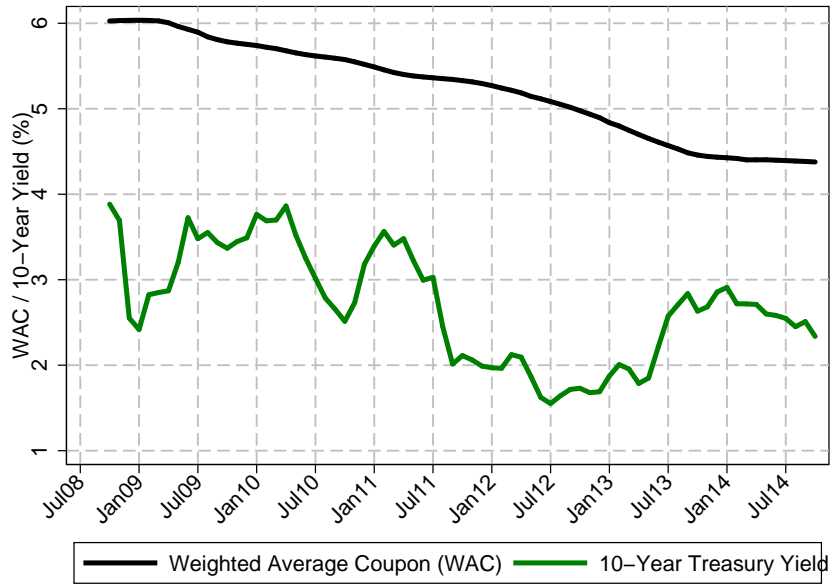


Figure 6: Estimated  $\phi$ ,  $\pi$ , and Covariates

Notes: As explained in Section 6,  $\phi$  is our baseline estimate of the price of intermediation and  $\pi$  is our alternative measure of intermediary margins where the 25 basis points of servicing cash flow are valued by a separate multiplier, rather than implicitly by MBS prices. Lower panels show monthly average  $\phi$  plotted against applications per business day and median processing time (from HMDA); implied interest rate volatility (MOVE); payroll per employee in real estate credit (from the BLS); and mortgage market concentration (from HMDA). *Source:* HMDA; Optimal Blue; J.P. Morgan Markets; authors' calculations.

### A. Weighted Average Coupon and Ten-Year Treasury Yield



### B. WAC–Ten-Year Yield and Applications

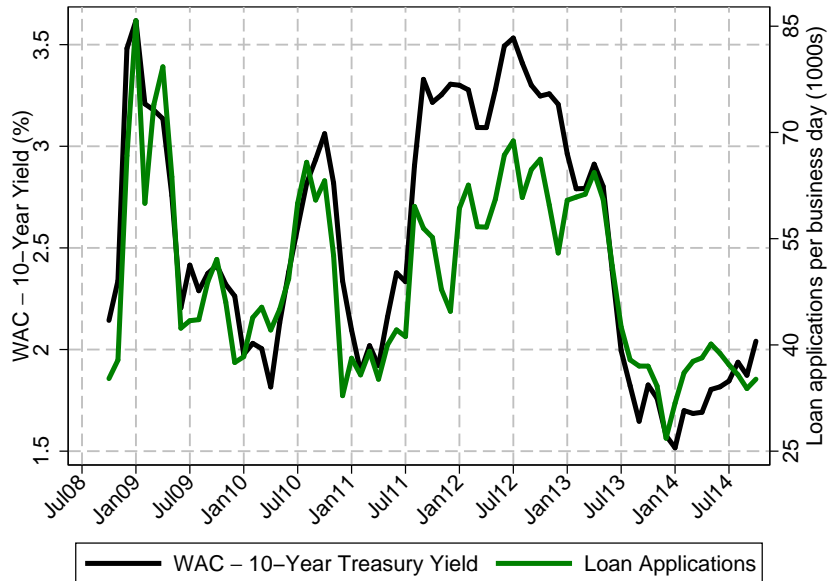


Figure 7: Gap between the Weighted Average Coupon and Ten-Year Treasury Yield, and Relation to Applications

Notes: Top panel shows the weighted average coupon (WAC) on outstanding conventional MBS and the ten-year constant maturity Treasury yield over our sample. Bottom panel shows the difference between the WAC and ten-year yield against the monthly number of new loan applications (divided by the number of business days). Sources: J.P. Morgan Markets, FRED, HMDA.

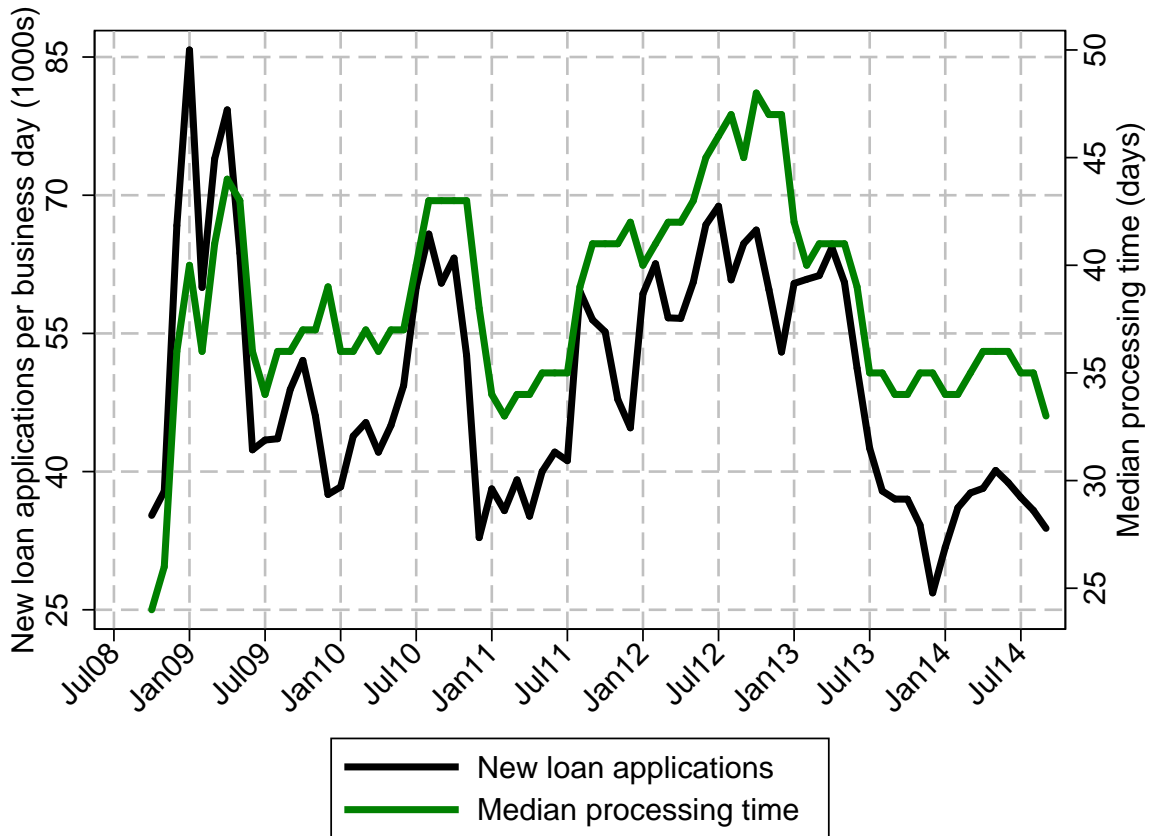
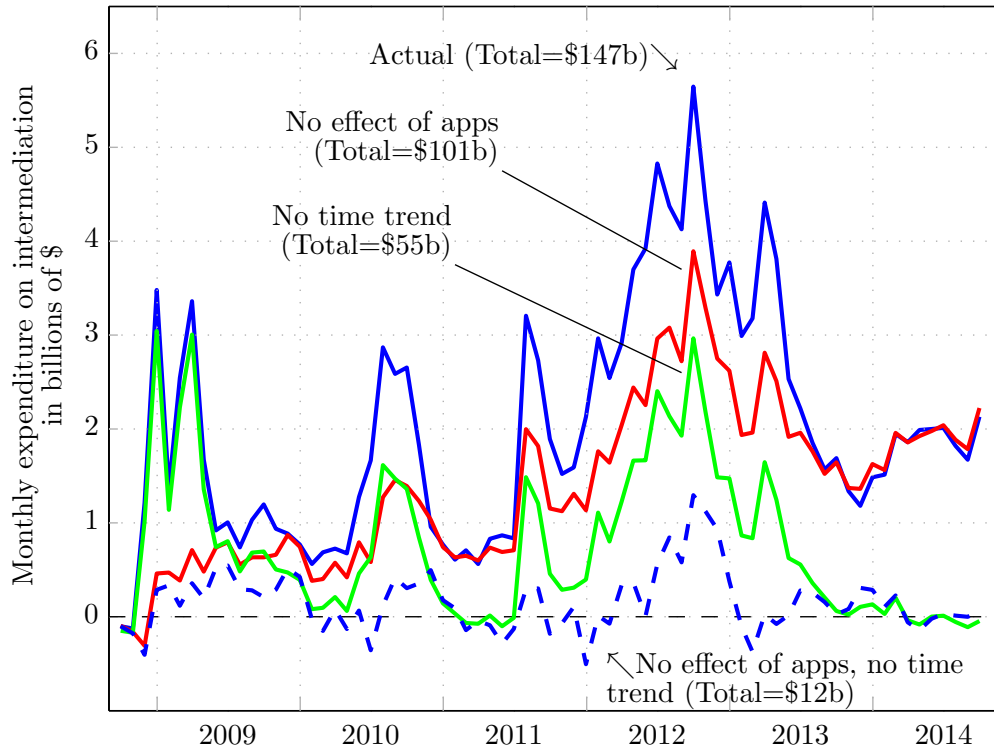


Figure 8: Loan Application Volume and Median Processing Time

Notes: Figure shows the evolution of the total number of new loan applications in a month (divided by the number of business days in that month) and the median processing time for loan applications submitted in the corresponding month. Processing time is measured as the number of days between the application date and the “action” date, when either the loan closes, or the application is denied or withdrawn. Source: HMDA.

### A. Expenditure on Intermediation with Counterfactuals



### B. Rate101 with Counterfactuals

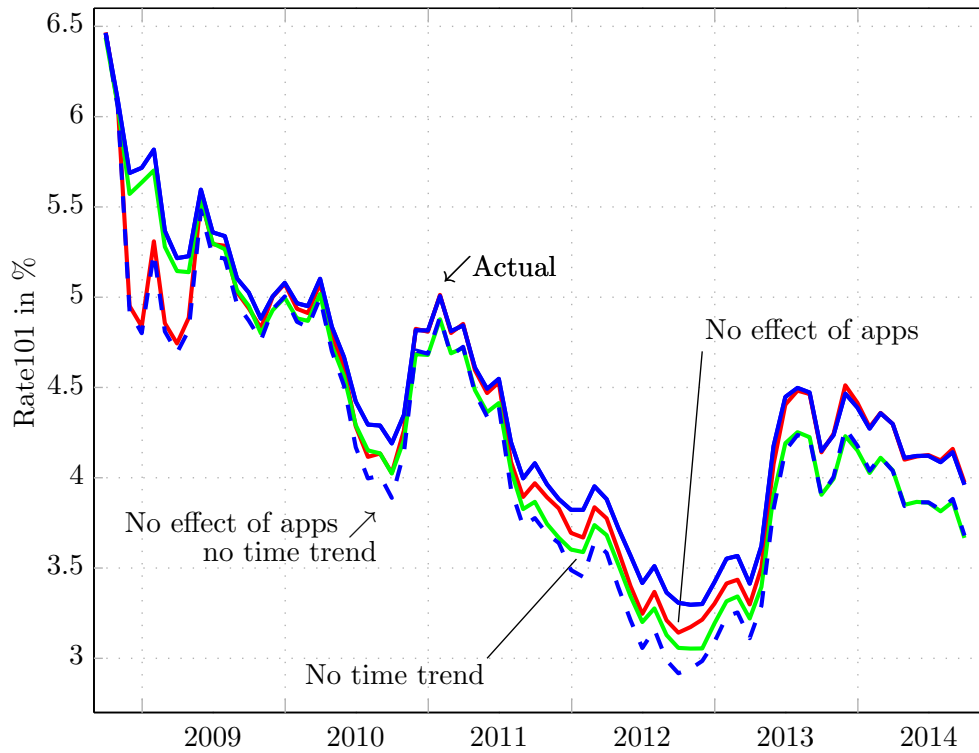


Figure 9: Expenditure on Intermediation and Rate101 Under Different Assumptions

Notes: Top panel shows our estimate of total household expenditure on intermediation over the sample period. We generate counterfactual paths using regression estimates from Table 4. Bottom panel shows actual Rate101 counterfactual rates under same assumed counterfactuals. See Section 7.1 and 7.2 for relevant discussion.

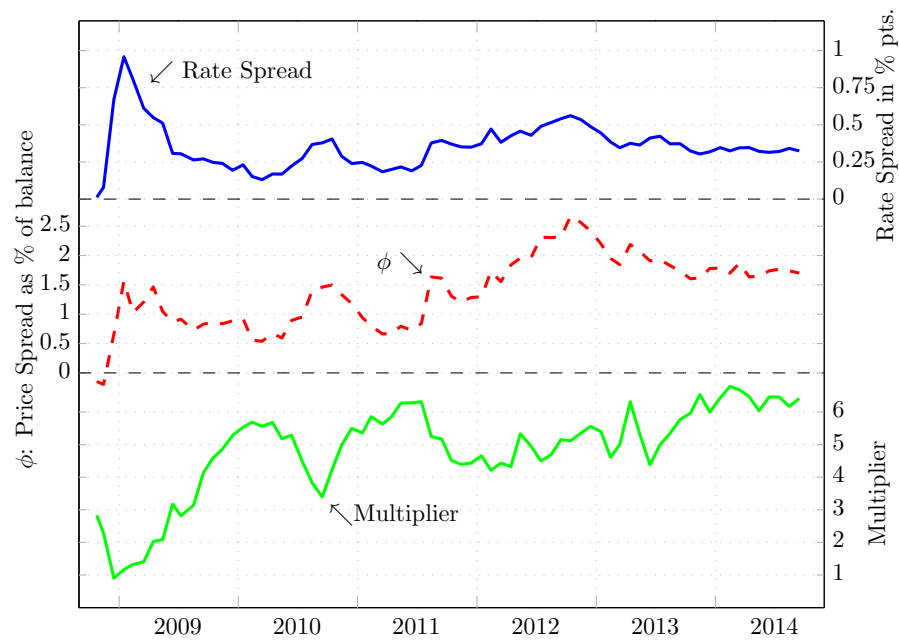


Figure 10: Primary-Secondary Spread Versus  $\phi$

Notes: As explained in Section 7.2,  $\phi$  equals the primary-secondary rate spread times the multiplier. To calculate the rate spread, we first use MBS prices to compute  $y^{101}$ , the note rate on a loan with a 1 percent rebate assuming that the borrower could borrow directly in the TBA market. To get the rate spread, we subtract  $y^{101}$  from  $Rate_{101}$ . The multiplier is computed using the MBS coupons trading immediately above and below 101.

# Online Appendix for “The Time-Varying Price of Financial Intermediation in the Mortgage Market”

Andreas Fuster, Stephanie H. Lo, and Paul S. Willen

## A Additional Methodology Details

### *Rate sheet data from Optimal Blue*

1. Over the entire sample period, there are 49 days missing, either due to missing back-ups/queries or due to obviously inconsistent data.
2. In some cases, a given LO will receive multiple rate sheets from the same lender. In those cases, we select the highest rebate for a given note rate.
3. When our data provides rate sheets from a given lender for multiple LOs, we average rebates across LOs for a given note rate.
4. While most rates are quoted in 1/8 of a percentage point, rate sheets sometimes have exceptions (e.g. 4.99 instead of 5). We round note rates within 2 basis points of 1/8 of a point to the nearest 1/8.
5. We omit dominated offers. For example, if a lender offers a 5.25 percent note rate with a rebate of 3.0 and a 5.125 percent note rate with a rebate of 3.5, we omit the former offer.

### *Aggregation across lenders*

1. For pass-through regressions (Sections 4 and 5), we first calculate *Rate101* lender-by-lender. We then calculate  $\Delta p_{YSP}$  lender-by-lender and took the median. To calculate  $\Delta p_{TBA}$ , we use the median *Rate101* across lenders. Other aggregation approaches yield more or less identical results.
2. To calculate  $\phi$  and  $\pi$  at monthly frequencies, we first calculate *Rate101* at a daily frequency as explained in the text meaning that the rebate was by construction equal to one for all lenders and  $p_{YSP} = 101$ . We then use *Rate101* to calculate  $p_{TBA}$  and subtract  $p_{YSP}$  to get  $\phi$ . We then average across days of the month to get a monthly value for  $\phi$ .

*HMDA data:* We exclude multifamily housing, junior liens, and preapproval requests that were denied or approved but not accepted. For any given day, we count new applications as those that were applied for on that day. To be included in the HMDA data for a given year, an application needs to be processed within that year. Applications submitted towards the end of year  $t$  are thus frequently included in the HMDA file for year  $t + 1$  only. Since we do not have the 2015 HMDA data yet, including November-December 2014 would lead us to understate application volumes for these months.

## B Pass-Through in Rate-Yield Space

In Table A.1, we report the results from pass-through regressions similar to the ones reported in column (2) of Table 3 in the main text. However, instead of analyzing changes in rebate/price space, we now run regressions of changes in the mortgage rate on changes in MBS yields. For mortgage rates, we use either our Rate101 (available daily) or the Freddie Mac Primary Mortgage Market Survey rate (available weekly). For MBS yields, in line with what researchers studying this market would typically do, we rely on the “current coupon” MBS yield (from J.P. Morgan). An increase in MBS prices corresponds to a decrease in MBS yields.

The first three columns of the table use Rate101. In column (1), we see that at a daily frequency, the estimated pass-through coefficients are almost identical to those in Table 3, although the  $R^2$  is lower, suggesting more noise in the rate/yield measures than in rebates/prices. The asymmetry discussed in the main text is also present for rates: an MBS yield increase is on average completely passed through to rates, while a yield decrease passes through by only 80 percent.

In column (2), we take the weekly averages of both rates and yields. The estimated coefficients stay almost identical, while the  $R^2$  increases somewhat. In column (3), we do the same calculation at the monthly frequency.<sup>1</sup> Perhaps surprisingly, pass-through appears to get weaker when taking monthly averages. This is likely due to the fact that, as emphasized in our main analysis, lower frequency shifts in the price of intermediation (due to changes in demand, or cost shifters) will be reflected in these longer-horizon pass-through estimates (while at the daily frequency these are more likely to remain unchanged). Furthermore, as discussed in Section 7.2, rate/yield changes also reflect changes in the valuation multiplier.

In the remaining columns of the table, we repeat the weekly and monthly regressions

---

<sup>1</sup>For both weekly and monthly averages, we only retain MBS yields for days on which we also have Rate101 (as noted in Appendix A, there are 49 days missing). Furthermore, we only retain months for which we have at least 18 daily observations.

using the Freddie Mac rate instead, over the same sample period for which we have Rate101 (columns 4 and 6) and also over a longer period starting in 1992 (columns 5 and 7).<sup>2</sup> We also control for changes in reported origination fees/points (though not doing so leaves the other coefficients almost unchanged). Column (4) shows that at a weekly frequency, the estimated pass-through is significantly weaker than with Rate101, suggesting that Rate101 may be a “cleaner” measure since it holds borrower/loan characteristics exactly constant. In contrast, the “noise” in the Freddie Mac rate could be correlated with changes in MBS yields, thereby somewhat attenuating the coefficients. At a monthly frequency, as shown in column (6), the Freddie Mac results move closer to those based on Rate101. Looking at the longer sample does not lead to material differences from the coefficients at a weekly frequency (column 5), though it does increase the estimated pass-through at a monthly frequency (column 7).

In sum, the finding of an asymmetric pass-through of changes in the MBS market value of a loan to the primary market is robust to looking at rates instead of rebates, to using different rate series, and to different horizons over which changes are computed. The pass-through does not increase when looking at longer horizons because of changes in the factors driving the price of intermediation and changes in valuation multipliers.

## C Other Measures of Intermediation Prices and Costs

In this section, we briefly discuss other methods of measuring the price of intermediation. First, we consider lenders’ reported production income. Production income equals the non-interest revenue generated by mortgage originations and sales, and thus should be correlated with  $\phi$ . We calculate production income in two ways. First, we use SEC filings for the “big four” banks (Bank of America, Citigroup, J.P. Morgan, and Wells Fargo), and sum the total income across them. For all banks except Citigroup, we use the line “Net gains on mortgage loan origination/sales activities” for mortgage banking activities from the quarterly 10-Q filings. For Citigroup, we use the line, “Net Servicing & Gain/(Loss) on Sale” for North America from their Quarterly Financial Data Supplement. We then divide this sum by the total volume of these banks’ mortgage originations (also from the 10-Qs). The line labeled “Production income at Big 4 Banks” in the top panel of Figure A.2 shows the resulting series, which has a correlation of 0.74 with our measure of the price of intermediation,  $\phi$ .

An alternative measure of production income comes from survey data taken from the Mortgage Bankers’ Association (MBA). The MBA also calculates the costs associated with

---

<sup>2</sup>Since the Freddie Mac rate reflects loan offers from Monday to Wednesday within a week (according to [http://www.freddiemac.com/pmms/pmms\\_faqs.html](http://www.freddiemac.com/pmms/pmms_faqs.html)), we also only use MBS yields from Monday to Wednesday.



originating a loan and the line labeled “Net Production income” is production income net of costs. Net production income has a correlation of 0.78 with  $\phi$  despite the fact that the MBA series does not trend upward (which may be due to the fact that costs are subtracted).<sup>3</sup>

In Section 6.1, we argued that increases in the cost of underwriting and servicing were driving up the overall price of intermediation. The MBA survey data provides some direct evidence along those lines. The second panel of Figure A.2 shows the number of underwriters required to originate 100 loans per month which, according to the MBA, rose from less than one in 2008 to nearly three in 2013. This increase provides one potential explanation for the time trend in  $\phi$ .<sup>4</sup> The third panel of Figure A.2 shows the cost of servicing a performing loan which, according to the MBA, increased from \$60 per year in 2008 to nearly \$160 in 2013. To be clear, the increase in servicing costs depicted in the figure are for performing loans and thus do not have to do with the added burdens of servicing delinquent loans which, according to the same survey, went from under \$500 in 2008 to more than \$2300 per year in 2013.<sup>5</sup> Fratantoni (2016) provides further discussion of the MBA data showing increases in costs in recent years.

## D Robustness of $\phi$ and $\pi$ , and a Longer Time Series

For the regressions of  $\phi$  and  $\pi$  in the paper, we use the monthly averages of each series. Figure A.1 shows the daily series of  $\phi$  and  $\pi$ . While there is a significant amount of daily variability in the price of intermediation, the graph demonstrates that aggregating on a monthly basis does not result in the loss of any important daily variation.

In the paper, we use a base case of loan and borrower characteristics, as shown in Table 1. Figure A.3 shows that these assumptions are innocuous since  $\phi$  hardly changes as the underlying loan and borrower characteristics change.

In Figure A.4, we present a historical time series of  $\phi$ . Since we do not have the Optimal Blue data before October 2008, we rely on the 30-year conventional fixed rate obtained from the Freddie Mac Primary Mortgage Market Survey series (as in the previous section), subtract the g-fees, and interpolate between MBS coupons to derive our series for  $\phi$ .<sup>6</sup> We

<sup>3</sup>For details, see <https://www.mba.org/news-research-and-resources/research-and-economics/single-family-research/mortgage-bankers-performance-reports-quarterly-and-annual>

<sup>4</sup>Data comes from <https://www.mba.org/Documents/Research/ChartoftheWeek%2007102015.pdf>.

<sup>5</sup>See <http://www.urban.org/sites/default/files/alfresco/publication-pdfs/2000607-Servicing-Costs-and-the-Rise-of-the-Squeaky-Clean-Loan.pdf> for details.

<sup>6</sup>A similar figure is presented as Chart 2 in Fuster et al. (2013). There is a difference in levels between the chart in Fuster et al. (2013) and our  $\phi$  because here we subtract 101 from  $p_{TBA}$  while in the other paper it was only 100. There are also small differences in assumed g-fees between the two calculations, but these only have minor effects on the evolution of the series.

also plot a version where we add the “origination fees and discount points” from the PMMS to  $\phi$ , since these should arguably be integrated in our price of intermediation measure when derived from a rate that may partly reflect changing rebates chosen by borrowers. The PMMS origination fees/points series declines steadily over 1998-2006, by about 60 basis points. It then partially recovers in 2008, remaining steady over the sample period of our main analysis.

The chart illustrates a few points that are important in the context of our analysis. First, the level of  $\phi$  was not unusually low at the beginning of our main sample period (shaded in the chart)—it had been steady at a very similar level over the previous 2.5 years. Second, there was significant variation in  $\phi$  in earlier years as well; in particular, an increase over mid-2001 to 2004, which coincides with a previous period of high loan application volumes (corresponding to the largest refinancing wave in U.S. history). Third,  $\phi$  has remained quite steady over the two years following the end of our sample period (in October 2014), except for an increase following the drop in rates after the U.K. referendum (“Brexit”) in late June 2016, which led to an uptick in refinancing. Overall, this evidence suggests that  $\phi$  has stabilized at its new, higher level.

## E Variation in Expected Lifespan of New Loans

Section 6.2 in the main text discusses the possibility that the increases in  $\phi$  and  $\pi$  that we document could be driven by a lengthening of the expected lifespan of newly originated loans. This could be the case if intermediaries require upfront compensation for per-period costs over the life of the loan—for instance the risk (non-zero in each period in which the loan remains active) that may be forced to repurchase the loan from the securitizing agency (e.g. Fannie Mae). We already presented some evidence that seems inconsistent with this hypothesis. Here, we additionally look directly at how measures of the expected lifespan of newly originated mortgages have evolved and how they correlate with  $\phi$ .

It is difficult to obtain a good measure of the expected time a new loan will remain open, but we can use two proxies obtained from the J.P. Morgan MBS model (consisting of a prepayment model and an interest rate model). One is the “weighted average life” (WAL), which is the expected average time until the mortgage principal is repaid (either by scheduled amortization or by unscheduled prepayments). This is calculated under a single assumed mortgage rate path, which is a shortcoming because the desired measure would take into account that rates vary in the future, which in turn affects prepayments. A measure that does take this into account is the “option-adjusted duration” (OAD), which is the weighted average time until mortgage cash flows are received (both coupon payments and principal

payments). Again, this is not a perfect measure for us, since we are not interested in the coupon payments for our purpose, but changes in the OAD should be a good proxy for changes in expected lifespan of a new loan.

We obtain time series of the WAL and the OAD corresponding to Rate101 based on coupon-level measures from J.P. Morgan.<sup>7</sup> Figure A.5 plots these against  $\phi$ . The chart suggests two things: first, the spikes in the price of intermediation do not seem to be associated with spikes in the expected lifespan of new loans; if anything, the opposite relationship holds. Second, there is no upward trend in the WAL or OAD since early 2010, while  $\phi$  and also to a lesser extent  $\pi$  have drifted upward over that time.

Thus, based on this evidence it appears unlikely that high- or low-frequency variation in the expected life of new loans was an important driver of the evolution of the price of intermediation over 2008–2014.

## F Derivation of the Relationship Between Price and Rate Spreads

Consider an intermediary that makes a loan with note rate  $r^n$  and pays the borrower  $p_{YSP}^n$ . Rather than selling the loan, the intermediary finances  $p_{YSP}^n$  in the secondary market by issuing debt. Denote  $y^n$  to be the implied coupon on this debt, defined by:

$$p_{YSP}^n = PV(\text{principal payments}(n)) + y^n \text{mult}_{MBS}^n, \quad (9)$$

where  $\text{mult}_{MBS}^n$  is a valuation multiple similar to the one defined in Section 2.3.1. Similarly, we can write that

$$p_{TBA}^n = PV(\text{principal payments}(n)) + r^n \text{mult}_{MBS}^n. \quad (10)$$

Equations (9) and (10) generate a relationship between the price of intermediation measured as the difference in prices and as the spread between rates in the primary market and the implied coupon in the secondary market:

$$\phi^n = p_{TBA}^n - p_{YSP}^n = (r^n - y^n) \text{mult}_{MBS}^n. \quad (11)$$

In other words, the one-time payment to the intermediary,  $\phi^n$  equals the present value of an interest strip that pays the difference between what the borrower pays to the intermediary

---

<sup>7</sup>To do so, we take the note rate that J.P. Morgan assumes for each MBS coupon, and directly interpolate between these rates and corresponding WALs/OADs to where our Rate101 would be.

$(r^n)$  and what the intermediary pays to investors  $(y^n)$ .

Rearranging equation (11) as

$$r^n - y^n = \frac{\phi^n}{\text{mult}_{MBS}^n}, \quad (12)$$

and taking differences over time shows that changes in rates reflect three factors:

$$r^n(t') - r^n(t) = \underbrace{y^n(t') - y^n(t)}_{\substack{(1) \Delta \text{Secondary} \\ \text{Market Cost of} \\ \text{Funds}}} + \underbrace{\frac{\phi^n(t') - \phi^n(t)}{\text{mult}_{MBS}^n(t')}}_{\substack{(2) \Delta \text{Price of} \\ \text{Intermediation}}} - \underbrace{(r^n(t) - y^n(t)) \frac{\text{mult}_{MBS}^n(t') - \text{mult}_{MBS}^n(t)}{\text{mult}_{MBS}^n(t')}}_{(3) \Delta \text{MBS Multiplier}}. \quad (13)$$

In contrast, changes in the primary market price only reflect two factors:

$$p_{YSP}^n(t') - p_{YSP}^n(t) = \underbrace{p_{TBA}^n(t') - p_{TBA}^n(t)}_{\substack{(1) \Delta \text{Secondary} \\ \text{Market Price}}} - \underbrace{(\phi(t') - \phi(t))}_{(2) \Delta \text{Price of} \\ \text{Intermediation}}. \quad (14)$$

	Rate101			Freddie Mac			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta y_{TBA,t}^-$	0.803*** (0.048)	0.810*** (0.076)	0.756*** (0.076)	0.539*** (0.079)	0.601*** (0.034)	0.620*** (0.032)	0.748*** (0.037)
$\Delta y_{TBA,t}^+$	0.976*** (0.042)	1.002*** (0.057)	0.849*** (0.106)	0.797*** (0.073)	0.799*** (0.028)	0.847*** (0.072)	0.922*** (0.029)
$\Delta$ Points				-0.086** (0.037)	-0.044** (0.021)	-0.125 (0.133)	-0.076 (0.046)
Constant	-0.004** (0.002)	-0.009** (0.004)	-0.018* (0.011)	-0.013*** (0.005)	-0.010*** (0.002)	-0.023*** (0.006)	-0.017*** (0.005)
Adj. R <sup>2</sup>	0.70	0.78	0.85	0.77	0.78	0.92	0.94
Observations	1431	309	69	315	1190	72	273
Frequency	Daily	Weekly	Monthly	Weekly	Weekly	Monthly	Monthly
Sample Period	200810.201410	200810.201410	200811.201410	200810.201410	199201.201410	200811.201410	199202.201410

Robust standard errors in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A.1: Pass-Through from MBS Current Coupon Yields to Mortgage Rates

Notes: Dependent variable is the change in the mortgage rate, measured either by “Rate101” in the Optimal Blue data (see main text for details) or by the 30-year fixed-rate mortgage rate reported in the Freddie Mac Primary Mortgage Market Survey (PMMS).  $y_{TBA}$  denotes the “current coupon” MBS yield (meaning the yield on an interpolated MBS coupon trading at par), obtained from J.P. Morgan. In columns (4)-(7), “Points” refers to origination fees and points (as percentage of the loan amount) reported in PMMS.

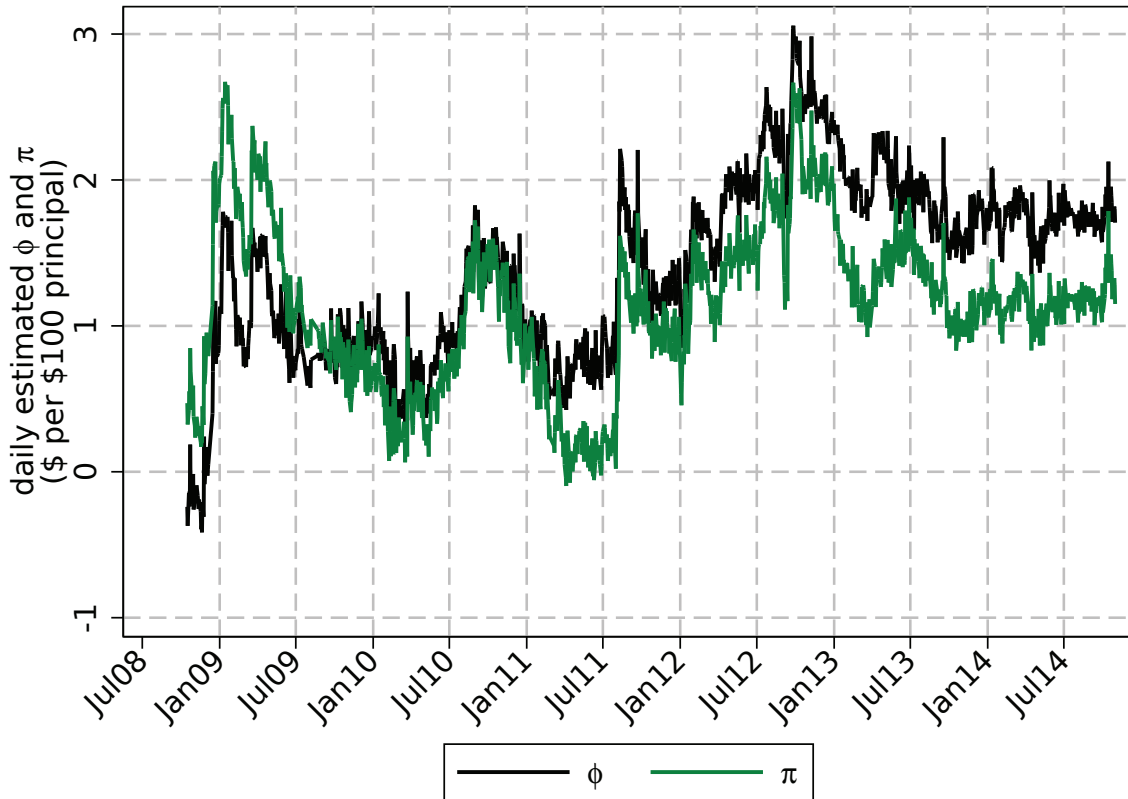
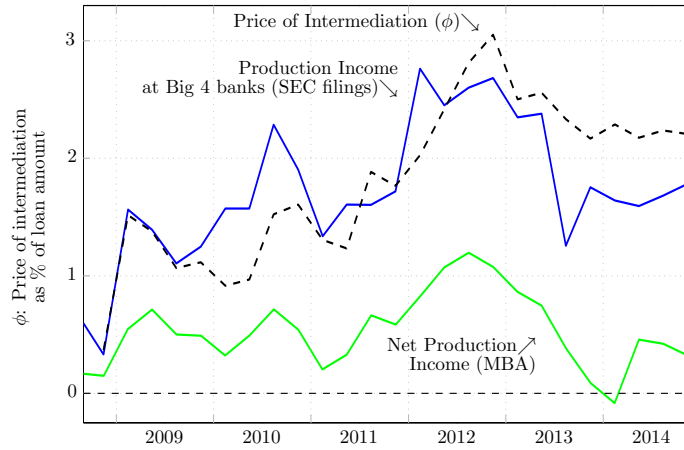


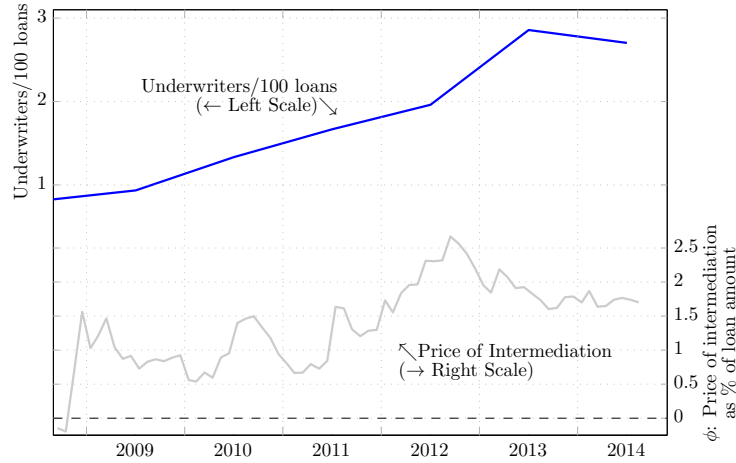
Figure A.1: Daily Series of  $\phi$  and  $\pi$

Notes: Daily version of the monthly series shown in Figure 6 of the main text.

### A. Production Income



### B. Mortgage Underwriting Productivity



### C. Cost of Servicing

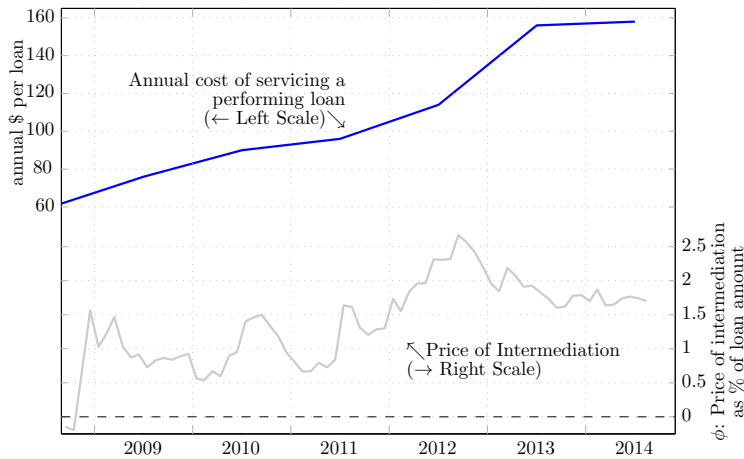


Figure A.2: Alternative Measures of the Price of Intermediation, and Direct Evidence on Evolution of Origination and Servicing Costs

Notes: See text in Appendix C for details and sources.

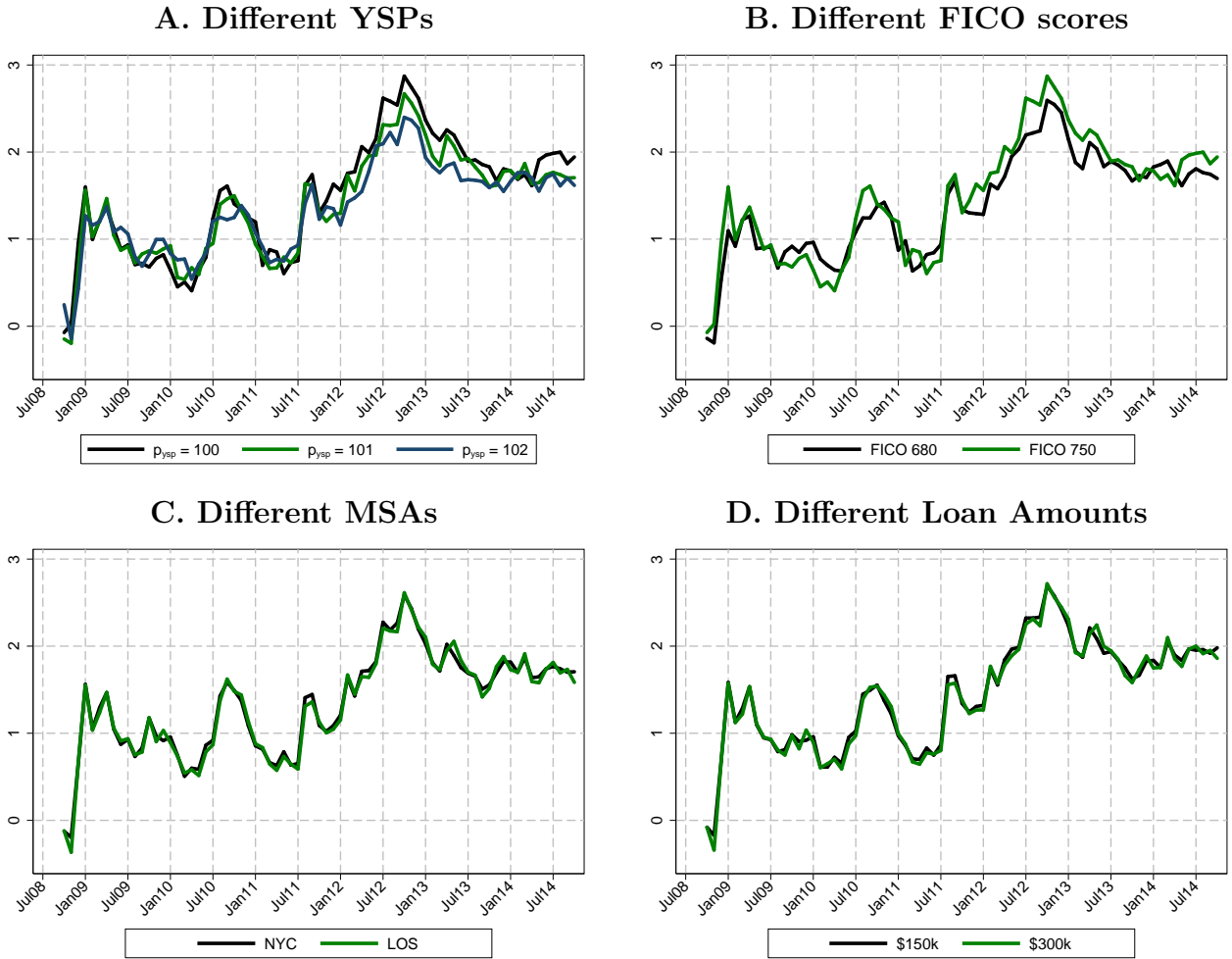


Figure A.3: Estimated  $\phi$  for Different Assumptions

Notes: Panels show  $\phi$  under different assumptions about the loan. Panel A compares  $\phi$  for three different mortgage note rates (our baseline assumption is  $p_{YSP} = 101$ ). Panel B compares  $\phi$  for a FICO credit score of 750 (baseline) with  $\phi$  for FICO of 680. (This is done at  $p_{YSP} = 100$ , since only few lenders offered 101 to FICO 680 borrowers in the early parts of our sample.) Panel C compares a loan in Los Angeles (baseline) to one in New York City. Panel D compares a loan for \$300,000 (baseline) to one for \$150,000.

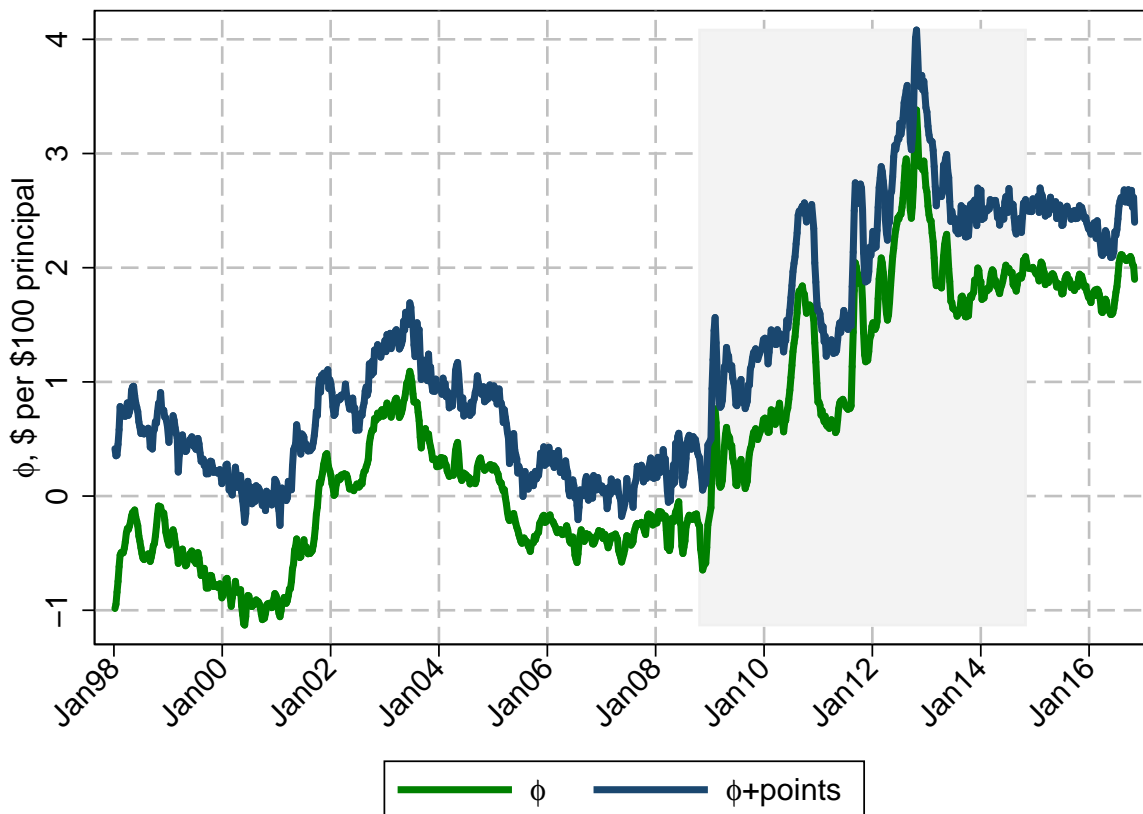


Figure A.4: Historical  $\phi$ , Rolling Four-Week Average

Notes: In contrast to the  $\phi$  series in the main paper, which is based on the Optimal Blue data, the historical  $\phi$  series is calculated using the 30-year fixed mortgage rate from the Freddie Mac Primary Mortgage Market Survey (PMMS), as explained in the text. In the “ $\phi + \text{points}$ ” series, “origination fees and discount points” from PMMS are added. Shaded area indicates time period covered in our main analysis.



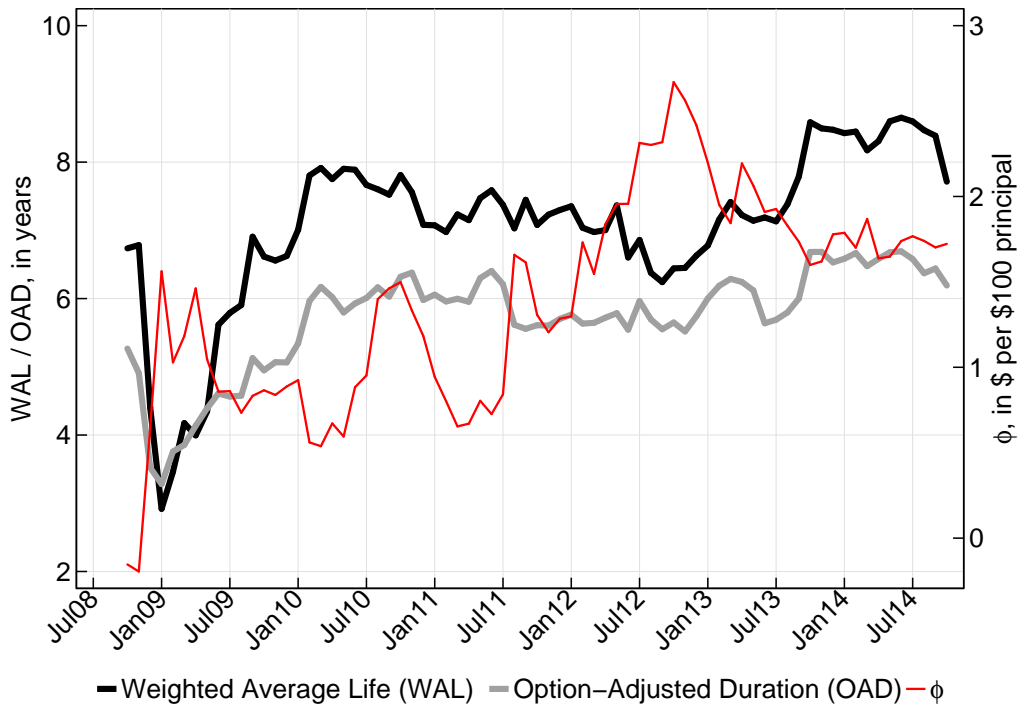


Figure A.5: Expected Life Span of New Mortgages and the Price of Intermediation  
 Notes: Weighted average life (WAL) and option-adjusted duration (OAD) are measures of the expected life span of newly originated loans, obtained from J.P.Morgan for different coupons and then interpolated to Rate101. See text in Appendix E for additional detail.