What You Don't Know Can't Pass Through: Consumer Beliefs and Pass-through Rates

Ryan M. Westphal*

March 9, 2024

Abstract

This paper shows that consumers' beliefs about firms' production costs are an important determinant of cost pass-through. By incorporating imperfect information about costs to a canonical model of consumer search, I show that changes in costs that consumers are aware of get passed through more completely than those they are unaware of. This model provides a first unified explanation of incomplete pass-through, over-shifting, and pass-through asymmetry. I test a novel prediction of this model using US mortgage data. I find that different components of the marginal cost of mortgage lending have different average pass-through rates. Widely known costs are passed through nearly completely while more obscure costs have much lower pass-through rates. This pattern is consistent with my model but cannot be explained by existing theory. I estimate that more precise information about lending costs would save small mortgage borrowers $289 million, annually.

*Westphal: Brandeis University and Research Department, Federal Reserve Bank of Boston, westphal@brandeis.edu, 480-776-9948, 415 South St, Waltham, MA 02453. I would like to thank Lucas Coffman, Michael Grubb, and Charles Murry for their invaluable feedback and advising. I would also like to thank Jason Abaluck, Susanto Basu, Ryan Chahrour, Christopher Conlon, James Dana, Michael Dickstein, Charles Hodgson, Maarten Janssen, David Laibson, Matthew Leisten, Julie Mortimer, Theodore Papageorgiou, Devesh Raval, Richard Sweeney, Danna Thomas, Andrew Usher, Matthijs Wildenbeest, seminar participants at the FTC, and participants at the 11th Workshop on Consumer Search and Switching Costs at NYU, 2022 WUSTL EGSC, and the 2023 IIOC for additional comments and discussions. Thank you to Paul Willen for numerous conversations about the mortgage market and data. All errors are my own. The views expressed in this paper are solely those of the author and not necessarily those of the Federal Reserve Bank of Boston nor the Federal Reserve System. **Keywords:** Pass-through, Beliefs, Consumer Search, Mortgages, Awareness.

1 Introduction

The extent to which producers' costs pass through to prices varies widely across markets and is important for understanding consumer welfare. This is especially true in an inflationary environment with rapidly changing costs. As fluctuating prices of materials and labor change the marginal cost of producing goods and services, pass-through rates allow economists and policy makers to study incidence and predict how much of the burden is borne by firms and how much is passed along to consumers through higher prices.

One popular piece of business advice on effectively transferring increased costs to prices is to explain the underlying cost increases to customers (Dholakia (2021); Heaslip (2022); Emmer (2022)). This is puzzling from the viewpoint of the existing theory on pass-through rates, which provides no scope for consumer awareness of firm costs to affect prices. In existing models of pass-through, consumers only affect pass-through rates via the shape of the demand curve, which is independent of the costs borne by firms.

In this paper, I study how consumers' awareness of firms' costs affects pass-through rates in markets with costly search. I show how a model of costly search with imperfect information about costs can explain a wide range of well-documented empirical facts about pass-through. This model provides a unified framework for explaining established empirical patterns of incomplete pass-through (pass-through less than one), over-shifting (pass-through greater than one), and asymmetric pass-through (greater pass-through of cost increases than decreases). The key result is that the extent to which firms' pass changes in costs to consumer prices increases with consumers' awareness of these cost changes.

The model captures the following intuition. If a consumer sees a high price following a well publicized increase in market wide production costs, the consumer will think that the high price is the result of the cost increase and they will not be able to find cheaper prices elsewhere. However, if a consumer sees a high price and has no idea that firms' costs have increased, they will think that other firms must still be offering lower prices and could do better by continuing to search. In equilibrium, firms increase their prices by less in the low information setting in order to deter search. Therefore, if consumers have imperfect information about the costs governing firms' pricing, then both *actual* costs and their *beliefs* about these costs will ultimately play a role in determining the distribution of prices.

Empirically testing the central prediction that pass-through is increasing in consumers' knowledge of cost changes presents a challenge. One needs to estimate pass-through in a setting where consumers' information about costs varies, but all of the other determinants of pass-through, such as the demand curve and firm conduct, are held constant. Any comparison of pass-through rates across markets or time will fail to hold demand constant and hence not be convincing. The solution is to focus on a novel prediction of my model: different cost components of the same product in the same market should have different pass-through rates. Specifically, cost components that consumers are aware of should have higher average pass-through rates than components they are unaware of. I therefore leverage unique institutional details of the US residential mortgage market to test these predictions of the model using confidential Home Mortgage Disclosure Act (HMDA) data.

I show that different components of lenders' costs get passed though at different rates. This is an ideal setting because there are many observable "components of costs" for lenders, and some of these components are commonly understood by consumers, while others are not. Specifically, borrowers are generally aware of the interest rate environment, which affects the payments made by mortgage backed securities investors to lenders, so-called TBA prices. Borrowers are likely unaware of special "payups" that MBS investors pay to lenders for certain types of low loan balance loans. I leverage discontinuities around loan size cutoffs for these specified pool payups to identify the pass-though of payups relative to the pass-through of the more transparent TBA price. I find that TBA prices are passed through at roughly unity, while specified pool payups are passed-though at roughly 0.4. These findings are consistent with the model and cannot be explained with existing theories of price competition. Finally, I show that this has meaningful consequences for consumer welfare: I estimate that more precise information about specified pool payups would save small mortgage borrowers $289 million, annually.

I establish the main result of the paper, that pass-through is increasing in consumers' awareness of cost changes, by developing a theoretical model of consumer search and studying it under different consumer information settings. First, I show that that if consumers have complete information about firms' costs, as in canonical search models (Burdett and Judd, 1983; Stahl, 1989, e.g.), costs pass-through to average transacted prices at a rate of one. This demonstrates that introducing search frictions alone is not enough to generate imperfect pass-through and establishes a benchmark to compare with models of incomplete information.

Second, I introduce consumers' imperfect information over firms' costs, similarly to Janssen, Pichler, and Weidenholzer (2011), and show that cost changes pass-through to average transacted prices at a rate proportional the the number of non-captive consumers (those who search for more than one price).[1] I decompose the cost into two parts: the part that consumers know (their expectation of the cost) and the part that they do not know (the difference between that expectation and the actual cost). I find that the part that consumers know about gets passed through more completely than the part that they do not. This can explain a range of pass-through rates, both above and below unity, holding consumer demand for the good fixed, but varying consumers' beliefs about costs. I also show that different components of marginal cost can have different pass-through rates, the prediction that motivates the empirical test of my model.

Finally, I establish the main result of the paper: that pass-through is increasing in consumers knowledge of costs. To do so, I develop a model of how consumers form their beliefs over firms' costs. This model nests Stahl (1989) and Janssen, Pichler, and Weidenholzer (2011) as two cases. I show that if consumers receive a series of noisy signals about firms' ever changing marginal costs, the average pass-through rate is decreasing in the variance of these signals. In other words, in environments where consumers can closely track the costs that firms face, pass-through is high. In markets where costs are opaque and unknown to consumers, pass-through is lower. This leads to incomplete pass-through on average.

The pass-through rate in this environment depends on signal realizations, so it is stochastic. It is greater than one with positive probability when signals are informative but imperfectly so. I show that when consumers have either perfect information about costs or no information about costs, the probability of a pass-through rate greater than one, or over-shifting as Pless and van Benthem (2019) call it, is zero. However, in between these extreme cases, the probability of over-shifting is strictly positive. Intuitively, over-shifting occurs when consumers beliefs about costs are very high relative to actual costs.

Finally, when firms can pay to credibly disclose their costs to consumers, firms publicise cost increases but not price decreases—leading to the well documented asymmetry in pass-through which Peltzman (2000) describes with "prices rise faster than they fall." As an example of such a scenario, consider restaurants printing signs saying "Due to the national shortage of avocados, guacamole prices have increased." I show that this ability to disclose

[1] While search behavior is an equilibrium object, it turns out this is equivalent to consumers with non-positive search costs. The fraction of non-captive consumers can be thought of as a measure of market power. When it is one, the market faces perfect competition. When it is zero, the Diamond (1971) paradox is realized, and all firms operate as monopolies in separate markets. In between these extreme cases, firms have some degree of market power. I assume that the share is non-zero throughout the paper.

costs generates an equilibrium in which costs above a certain threshold are disclosed, while costs below it are not. In this version of the model, costs above the threshold have a pass-through rate greater than or equal to unity, while costs below this value have a pass-through rate less than one.

I use the US residential mortgage market as a case study for the novel prediction that pass-through varies with consumers' information about firms' costs. The US mortgage market has three key features that make it ideal for studying this issue. First, it is well documented that borrowers face search frictions (Allen, Clark, and Houde (2014), Bhutta, Fuster, and Hizmo (2020))[2]. Very few borrowers receive more than a handful of interest rate quotes and almost half receive only one (Alexandrov and Koulayev (2018)). There is considerable price dispersion in interest rates paid even for identical borrowers on the same day (Bhutta, Fuster, and Hizmo (2020)).

Second, from a practical point of view, the mortgage market has incredibly detailed administrative data which is essential to accurately measure pass-through rates. I use a confidential version of the HMDA database linked to mortgage backed securities (MBS) data, that allows me to precisely observe components of cost that would not be possible with any publicly available data.

Third, and most importantly, mortgage pass-through rates can be measured separately for different cost components to test the prediction that, all else equal, pass-through rates increase with consumers' knowledge of cost changes. The marginal cost of lending can be broken down into multiple components that have varying levels of visibility to borrowers. This variation in consumers' knowledge about the costs makes the market the perfect laboratory to observe this mechanism. All of these components of cost impact firms' bottom line in the same way, but the pass-through rates on the components are quite different. Previous studies have documented a number of factors that cause pass-through to vary across markets. Weyl and Fabinger (2013) show that the curvature of the demand and supply functions, as well as firm conduct all determine pass-through. Muehlegger and Sweeney (2021) demonstrate that whether cost shocks are market-wide or firm specific is an important factor. In this setting, all components of cost are market-wide and are part of the

[2]While large search frictions in the mortgage market make it an ideal case study for the mechanism, it is far from the only market in which consumers must search to discover prices. Kaplan and Menzio (2015), for example, find that price dispersion characteristic of consumer search can be found in a wide range of markets for consumer goods. Furthermore, search frictions are prevalent in other large, high-stakes markets such as the market for automobiles (Moraga-González, Sándor, and Wildenbeest (2023)). This model potentially has implications for pass-through in any such market across the economy.

same product; therefore, the demand curves and firm conduct are held fixed. If different components of cost have different pass-through rates in the same market, with the same firms, and the same borrowers, then we can plausibly attribute the difference in prices to the difference in consumers' information about each cost.

In this setting, I show that two components of the cost of lending are passed through at dramatically different rates. One of these costs, the cost of funding through the MBS market, closely tracks a number of widely publicized and tracked measures of interest rates. Mortgage borrowers are generally well aware of this component of the cost of lending and previous studies have found that it is passed through almost completely. The other determinant of cost, the specified pool payup, is idiosyncratic to the borrower, more opaque, and less well publicized. Consumers have less precise knowledge of the payup component of the cost and I find that it is passed through at a far lower rate. These findings are consistent with a model in which consumers have more precise expectations on the MBS yield than on payups. I find that the difference in pass-through rates between these components is robust to a wide range of assumptions about the total price that borrowers pay.

While these estimates of pass-through are primarily used as a test of the theoretical model, they can also be used to estimate the consumer welfare gains from specified pool payups. Furthermore, they allow me to compare how these welfare gains would change under a counterfactual consumer information regime. Specifically, I compare current consumer welfare gains from specified pool payups to what these gains could be if consumers had equally precise information about payups and general MBS yields. I find that if the average specified pool eligible consumer had more precise information, they would save between \$670 and \$905 in upfront payments. I estimate that if small loan size borrowers whose loans are guaranteed by Fannie Mae had more precise information about specified pool payups, they would save a total of \$289 million, annually.

The rest of the paper is structured as follows. Section 2 places this paper in the context of the related existing literature. Section 3 develops a theoretical model of pass-through with search and imperfect information. Section 4 gives an overview of the US residential mortgage market. Section 5 discusses the empirical strategy used to test the predictions of the theoretical model and presents the empirical results. Section 6 discusses implications of the model and concludes.

2 Literature Review

The literature measuring the pass-through rates of marginal costs is extensive. To the best of my knowledge, this is the first paper to measure the pass-through rate of multiple separate components of marginal cost for the same product and demonstrate that they are different — a result that is not explained by existing models of equilibrium pass-through. Economists have studied the pass-through to prices of costs from material inputs such as oil (Borenstein, Cameron, and Gilbert (1997)) and coffee beans (Nakamura and Zerom (2010)); taxes such as excise taxes on cigarettes (Sumner (1981), Bulow and Pfleiderer (1983) or fuel (Marion and Muehlegger (2011)); exchange rate fluctuations (Goldberg and Knetter (1997)); monetary policy (Fuster, Lo, and Willen (2017)); and many others. Pass-through rates have been used indirectly to study welfare analysis (Chetty (2009)), demand elasticities (Miller, Remer, and Sheu (2013), and merger implications (Jaffe and Weyl (2013), Miller, Remer, Ryan, and Sheu (2016)), among other things.

Weyl and Fabinger (2013) provide extensive analysis on how market "conduct" determines pass-through and ultimately incidence. In a model that nests perfect competition, monopoly pricing, and many commonly used models of imperfect competition, they demonstrate how to map the market structure into a sufficient statistic of the degree of competition in a market which then can be used in the calculation of pass-through rates and incidence. While the work of Weyl and Fabinger (2013) has been useful in studying the effect of market structure on welfare, one notable limitation is that it only focuses on markets with complete information. Critically, this rules out markets in which consumers are not aware of the prices offered by every firm in the market and must perform costly search to learn these prices. The model presented in this paper allows for incomplete information about prices, giving rise to a beliefs channel in pass-through.

This paper also contributes to theories of pricing in the presence of search frictions. Seminal models of equilibrium under search such as Burdett and Judd (1983) and Stahl (1989), generate price dispersion due to consumers' incomplete information about prices in the market. Benabou and Gertner (1993) and Fishman (1996) introduce consumer uncertainty over production costs in which consumers are unable to differentiate between industry-wide and firm specific cost shocks. Dana Jr (1994) presents a theory of consumer learning about an industry-wide cost. The model in this paper is most closely related to Janssen, Pichler, and Weidenholzer (2011). Their paper features consumers learning about an industry's stochastically drawn costs through sequential search. They relate the distribution of prices to both actual costs and consumers' expectations of costs. This paper expands upon their analysis by placing structure on how consumers initially form beliefs over costs for a given cost shock. This framework allows me to study how the precision of consumers' knowledge of the market state impacts pass-through of costs. This produces (and empirically tests) novel predictions that different components of cost for the same product can have different pass-through and that these pass-through rates are increasing in the precision of consumers' information. [3]

This paper is also related to the literature on awareness and salience. Chetty, Looney, and Kroft (2009) demonstrates that consumers underreact to taxes that are not salient. Kroft, Laliberté, Leal-Vizcaíno, and Notowidigdo (2023) model the interaction between tax salience and market power and empirically show that the incidence of sales taxes on many retail goods falls primarily on consumers. Busse, Silva-Risso, and Zettelmeyer (2006) shows that, in the automobile industry, well-publicized consumer rebates get passed through to consumers more completely than dealer discount promotions. The model presented here provides a distinct new channel through which tax salience may break the equivalence of statutory incidence of taxes. If consumers are unaware of a tax placed on producers, the tax may be passed through at a lower rate than a similar tax added to the price that consumers see. As in the above papers, my model predicts that consumers' information about taxes determines how much of the tax burden is borne by consumers.

Similarly, papers such as Bordalo, Gennaioli, and Shleifer (2013) and Bordalo, Gennaioli, and Shleifer (2015) model how consumers' awareness of different utility relevant components of the product (e.g. price or quality) affect consumer choice and firms' strategic decisions. Within the literature on financial products, there are many examples of consumers being unaware of payoff relevant elements of the product. Anagol and Kim (2012) finds that Indian mutual fund investors are inattentive to shrouded upfront fees. Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015) shows that limits on certain types of credit card fees were not undone by increases in interest rates, as interest rates are salient to consumers and these fees were not. Liu (2019) and Benetton, Gavazza, and Surico (2021) both study the UK mortgage market and find that borrowers are less sensitive to the mortgage contract's upfront fees than they are to interest rates and that lenders strategically respond

[3] Other papers such as Yang and Ye (2008) and Tappata (2009) show how non-sequential models of search can generate asymmetric pass-through. Cabral and Gilbukh (2020) studies asymmetric pass-through in a dynamic model of search where consumers only learn through prices. Janssen and Shelegia (2020) studies pass-through in a market with consumer uncertainty over costs but focuses on the specific informational setting of vertical industries.

to this difference in sensitivity. While this paper also studies consumers' awareness, it differs in a critical respect. All of the aforementioned papers focus on salience, or consumers' awareness, of payoff relevant aspects of a product. In each of the cited papers on financial products, consumers have less information about some component of the *price* they are paying. This paper studies how consumers' awareness of *firms' production costs*, a non-payoff relevant aspect, affects prices. Importantly, a model in which consumers have less info about some components of *price* than others, does not predict that different components of *firms' costs* will be passed through at different rates. This is, therefore, a novel mechanism through which consumers' awareness affects the pass-through rate of costs. I show that consumers' incomplete information about firms' costs can affect pass-through even when consumers fully understand and internalize all aspects of a product's price and its utility relevant attributes.

This paper relates to research in macroeconomics on the relationship between inflation expectations and realized price levels. In the textbook New Keynesian model, expectations over future inflation primarily affect the current price level through two channels (Weber, D'Acunto, Gorodnichenko, and Coibion (2022)). First, if firms face adjustment costs when changing prices, they must set their prices to maximize expected profit for the entire period that their prices are fixed. More closely related to this paper is the effect of consumer expectations of inflation on pricing. Theoretically, if consumers expect higher prices in the future, they should make durable goods purchases now while prices are lower. Research has shown that consumers do not form accurate expectations over future inflation and that these expectations are biased by personal experiences such as their own grocery shopping (D'Acunto, Malmendier, Ospina, and Weber (2021)). Most of this research describes the relationship between agents' expectations for *future* inflation and current prices. This paper illustrates a channel through which uncertainty over *current* inflation can affect current prices.[4]

[4]Within the macroeconomics literature there is also a strand of research on what authors call "customer markets," as in Rotemberg (2005). In these models, consumers react negatively to prices they perceive as unfair by not making repeated purchases with the firm. This dynamic consideration can generate price rigidity and lower pass-through of cost changes (Nakamura and Steinsson (2011)). Firms in my model make an entirely static pricing decision, but the intuition underlying the models is similar. Similarly, Eyster, Madarász, and Michaillat (2021) show that if consumers have preferences for fairness and update their beliefs about costs sub-proportionally, pass-through is incomplete.

3 Model

I examine a model in which N firms sell a homogeneous, indivisible good. A mass of consumers, have unit demand for this good at a value of v dollars, which is common knowledge to all agents. The market is characterized by search frictions similar to those found in Stahl (1989). Measure $1-\mu$ consumers search sequentially and must pay a search cost of s in order to learn a single price drawn randomly from the remaining firms. The other measure μ consumers are "shoppers," have a search cost of zero, and can costlessly view the prices offered by all of the firms in the market. Firms only make a single price offer to any given consumer.

To sell to a customer, a firm must pay a (common across firms) marginal cost of production C which is the sum of multiple components (indexed by j); so $C = \sum_{j=1}^{J} c_j$. I write the vector of these costs as $\boldsymbol{c}$. Each of these costs is drawn from a joint distribution function of costs and a signal x_j, $G_j(c_j, x_j)$ before sellers offer prices to consumers. Consumers do not directly observe any c_j.

The second random variable in this joint distribution is a signal x_j that both firms and consumers see. Normalize x_j such that, conditional on c_j, $E[x_j|c_j] = c_j$. Each cost component can then be rewritten as, $c_j = x_j + \epsilon_j$ where ϵ_j is, by construction, mean zero noise. In this framework, x_j represents the consumers' expectation of c_j and ϵ_j is the difference between reality and expectation. All consumers and firms observe the same vector of signals $\boldsymbol{x}$. Denote the distribution of ϵ_j conditional on x_j as $H_j(\epsilon_j|x_j)$.

I propose a *reservation price equilibrium* in the spirit of Stahl (1989). In a reservation price equilibrium, the consumers' optimal strategy is straightforward. Each consumer has a reservation price r and if they search and receive a price that is below r, they buy the good at that price. If they only receive prices that are higher than r, they search again. Below, I seek to characterize this reservation price and the distribution of prices offered by firms, $F(p)$. These values differ from the equilibrium values found by Stahl (1989) and later researchers in that r is a function of the signals received by the consumer. For that reason I write r as a function of the vector $\boldsymbol{x}$, $r(\boldsymbol{x})$.

The key object of interest is the pass-through rate of costs to prices. Because this environment is characterized by price dispersion in equilibrium, I focus on pass-through to the mean transacted price. I denote this price $\bar{p}(\boldsymbol{c}, \boldsymbol{x})$. I denote the pass-through rate of each component of cost as $\rho_j = \frac{\partial \bar{p}(\boldsymbol{c},\boldsymbol{x})}{\partial c_j}$.

In this framework, the pass-through rate depends on how consumers beliefs about costs respond to actual changes in costs. To see this, we can use the decomposition of cost from above, $c_j = x_j + \epsilon_j$, to rewrite the definition of pass-through as a total derivative.

$$\rho_j = \frac{\partial \bar{p}(c,x)}{\partial c_j} = \frac{\partial x_j}{\partial c_j} * \frac{\partial \bar{p}}{\partial x_j} + \frac{\partial \epsilon_j}{\partial c_j} * \frac{\partial \bar{p}}{\partial \epsilon_j} \tag{1}$$

Here one can see that ρ_j is a function of how consumers' information (x_j and ϵ_j) affects prices as well as how actual costs (c_j) affect consumers' information. This model therefore has two components: a model of the interaction of consumers and firms in this market and a model of consumer belief formation over costs. The model of consumer and firm interaction in a market with search costs and imperfect information determines how prices change as consumers' beliefs change ($\frac{\partial \bar{p}}{\partial x_j}$ and $\frac{\partial \bar{p}}{\partial \epsilon_j}$). The model of consumer belief formation then dictates how beliefs change as costs change ($\frac{\partial x_j}{\partial c_j}$ and $\frac{\partial \epsilon_j}{\partial c_j}$).

In Subsection 3.1, I consider the familiar case in which consumers have perfect information about costs. In this example, $x_j = c_j$, so we only need to consider one of the above partial derivatives. Proposition 1 tells us that pass-though is complete in this special case, establishing a benchmark to compare to cases of imperfect information. Then, in Subsection 3.2, I pin down $\frac{\partial \bar{p}}{\partial x_j}$ and $\frac{\partial \bar{p}}{\partial \epsilon_j}$ for a general class of beliefs. Proposition 2 states that the parts of firms' costs that consumers are aware of get passed through more completely than those the consumer is unaware of. Here, I also motivate this paper's empirical tests with Corollary 2.1. Having defined how prices react to beliefs, one can use any model of consumer belief formation to close the model and fully define pass-through. In Subsection 3.3, I suggest one such model of belief formation, in which Bayesian consumers form beliefs rationally in response to costs that follow a Gaussian process. This model gives us Proposition 3, which predicts that pass-through is increasing in the precision of consumers' information about costs. Finally, in Subsection 3.4, I extend the model to allow firms to reveal their costs to consumers. This extension generates asymmetric pass-through which is stated in Proposition 4.

3.1 Perfect Signals - Benchmark Pass-through

As a benchmark, consider the case where consumers perfectly observe firms' costs. In this case, $H(\boldsymbol{\epsilon}|\boldsymbol{x})$ is a degenerate distribution ($\epsilon_1 = \epsilon_2 = 0$ with probability 1) and thus we can write any reference to $\boldsymbol{x}$ as $\boldsymbol{c}$. This case corresponds to the model in Stahl (1989). A symmetric Nash equilibrium of the model is characterized by all firms mixing according to a price distribution $F(p)$ with support $p \in [\underline{p}, \bar{p}]$ and a consumer reservation price r, such that all prices in the support of F are in the best response correspondences of the firms and r is chosen optimally with respect to F. Given these assumptions, we can state the following:

Lemma 1 *When consumers perfectly observe firms' costs, and $\mu > 0$, a reservation price equilibrium exists.*

Proof: Stahl (1989).

It is useful to derive this equilibrium in order to understand the decisions facing consumers and firms. This is done in two parts. First, I derive the optimal firm pricing strategy, conditional on the existence of a reservation price equilibrium with reservation price r. Second, I derive the optimal reservation price strategy for the consumer, conditional on firm's pricing.

3.1.1 Equilibrium Firm Pricing

Assume that there exists a reservation price equilibrium with reservation price r. Firms' profit function from playing a price p is:

$$\pi(p) = \begin{cases} \left(\frac{1-\mu}{N} + \mu(1-F(p))^{N-1}\right)(p - C) & p \leq r \\ 0 & p > r \end{cases}. \tag{2}$$

If a firm plays a price greater than the consumer's reservation price, the consumer will search again, and the firm will not sell to this consumer (we will see that this holds in equilibrium). For prices less than r, the firm receives it's markup, $(p - C)$, multiplied by the share of consumers it receives. This share is the sum of the probability of seeing a captive consumer who has one price quote and the probability of seeing a shopper consumer who has N price quotes multiplied by the probability of being the lowest price of those N quotes. To ease notation, all further references to the profit function will only consider $p \leq r$ and should be thought of as the profit from playing any price less than or equal to the reservation price.

For this equilibrium to hold, firms must be indifferent between all prices in $[\underline{p}, r]$. Then, noting that $F(r) = 1$, profits from playing any price in this interval must be equal to profits from playing r which are defined by:

$$\pi(r) = \frac{1-\mu}{N}\,(r - C)\,. \tag{3}$$

The lower limit of the equilibrium price distribution, the point at which $F(\underline{p}) = 0$, can be solved for by equating the profit function (Equation 2) and profits at $p = r$ (Equation 3) and setting $F(\underline{p}) = 0$. Doing so gives us the lower limit of the support of F, $\underline{p}$. This is the highest price at which a firm will sell to all customers who receive a price quote from the firm. It is obvious that selecting a price $p' < \underline{p}$ is strictly dominated by playing $\underline{p}$ because p' will generate the same market share but with lower margins.

We can then solve for the price distribution F by setting (2) equal to (3). Thus, each firm will select a price in the interval $[\underline{p}, r]$ according to the distribution F defined by:

$$F(p) = 1 - \left(\frac{(1-\mu)(r-p)}{N\mu(p-C)}\right)^{\frac{1}{N-1}}. \tag{4}$$

We are generally interested in the pass-through to average prices, which requires us to integrate over the density of prices which I will denote $f(p)$.

3.1.2 Equilibrium Reservation Price

For the remainder of this section, to simplify the algebra, I focus on the $N = 2$ firm case.[5] Given the firms' pricing strategy, we need to find the reservation price for the non-shopper consumer (shoppers, of course, simply purchase the good at the lowest price they see). The non-shopper consumer compares utility from buying at the lowest price they have seen so far, to the expected utility from searching an additional time. The reservation price is pinned down by the price quote that makes these two values equal. It is defined implicitly by:

$$r = s + \int_{\underline{p}}^{r} [f(p)p]dp \tag{5}$$

Plugging in the density and $\underline{p}$ we find that r is:

$$r = C + \frac{s}{A(\mu)} \tag{6}$$

[5] The proof of the proposition in Online Appendix A does not depend on this restriction.

Where $A(\mu) = 1 - (\frac{1-\mu}{2\mu}) \log(1 + \frac{2\mu}{1-\mu})$. A is increasing in μ, meaning consumers' reservation prices are decreasing in μ. As Stahl (1989) notes, $\mu = 0$ corresponds to the Diamond (1971) monopoly pricing result, $\mu = 1$ corresponds to Bertrand marginal cost pricing, and values in between obtain intermediate results. We can therefore think of μ as a notion of "competitiveness" in the market. Naturally, consumers' reservation prices are decreasing in how competitive the market is. Reservation prices are also increasing in the search costs faced by consumers. Note that reservation prices are increasing one to one with the marginal cost. Search costs and μ both influence reservation prices but they do not change the relationship between marginal cost and the reservation price.

3.1.3 Pass-through

Using these equilibrium values, the following result can be obtained:

Proposition 1 *When consumers perfectly observe firms' costs, and $\mu > 0$, a reservation price equilibrium exists in which the pass-through rate of all marginal costs to average transacted price, ρ, is 1.*

Given consumers' reservation prices, we can calculate the distribution of transacted prices, the density of transacted prices, and ultimately the average transacted price ($\bar{p}$):

$$\bar{p} = C + \frac{(1-\mu)s}{A(\mu)} \tag{7}$$

Clearly the pass-through rate is 1. Levels of prices are affected by search costs and search technologies, but pass-through is not. This demonstrates that introducing search frictions (in this particular manner, at least) is not enough to introduce imperfect pass-through.[6] Furthermore, all costs have the same pass-through rate of 1. Each component of cost enters the expression only through C.

3.2 Imperfect Signals - Consumer Beliefs Affect Pass-through

Now assume that consumers do not perfectly observe firm's costs. Instead, they know the marginal distribution of costs conditional on the signal, $\boldsymbol{x}$, that they received. As before, the

[6]This differs from the results of Montag, Sagimuldina, and Schnitzer (2021) because they hold fixed the reservation price of the consumer. If one allows the reservation price to be an endogenous value pinned down by indifference towards search, rather than their maximum willingness to pay for the good, then markups are not a function of marginal cost, and thus pass-through is one.

residual part of cost unknown to consumers is captured by $\sum_j \epsilon_j$. When signals are imperfect, the distribution of $\sum_j \epsilon_j$ must be bounded from above in order to rule out improbable equilibria. Formally, we must assume:

Assumption 1 *There exists a value $\bar{\epsilon} \leq \frac{s}{A(\mu)} + E[\epsilon|p = r^*]$ such that $h(\epsilon) > 0 \implies \epsilon \leq \bar{\epsilon}$.*

To simplify analysis, we will also assume that the distribution of each ϵ_j does not depend on the expectation of cost. In other words, the belief distribution of costs shifts rather than reshapes as costs shift.

Assumption 2 *For each j, $H_j(\epsilon_j|x_j) = H_j(\epsilon_j|x'_j) \equiv H_j(\epsilon_j), \forall x_j, x'_j$ and $\epsilon_j < \bar{\epsilon}$.*

An example of consumer beliefs that meet both of these conditions would be that the total ϵ is uniformly distributed between zero plus or minus the markup charged by firms in the perfect information case. This model is closely related to that of Janssen, Pichler, and Weidenholzer (2011) but introduces a general class of signals about costs that are amenable to the analysis of pass-through. Finally, we will assume that there are only two firms, for simplicity.

Proposition 2 *When consumers imperfectly observe firms costs, $\mu > 0$, $N = 2$, and the distribution of the ϵ terms meets Assumption 1, a reservation price equilibrium exists. If Assumption 2 is also met, the pass-through rate of costs captured by consumers beliefs (x) is 1 and the pass-through rate on costs that are not (ϵ) is μ.*

To see this, assume again that there exists a reservation price equilibrium. Given a reservation price, everything from Subsection 3.1.1 holds. The firm's profit function is the same as before, and all characterizations of the distribution of offered prices and its support will be the same as functions of a reservation price. However, the consumer now faces a slightly different problem in setting r. The consumers' search problem now involves updating based on multiple pieces of information: the explicit cost signals and the prices received. The consumer sets their reservation price so that after receiving a price of r, their beliefs are such that they are indifferent between searching and not searching. Consumers with more price quotes of r are less likely to search than consumers with fewer price quotes of r. This means that we can pin down the search condition by the price that makes consumers with one price quote indifferent. Consumers with more price quotes may end up with a higher reservation price, but firms will never set a price above r because, in those cases, they are always in competition with at least one other firm, and in equilibrium, would lose with certainty. I therefore start off by looking at the search indifference condition for a non-shopper consumer.[7]

$$r = s + \underbrace{\int_{\epsilon_1}\int_{\epsilon_2} h_1(\epsilon_1, p = r) * h_2(\epsilon_2, p = r)}_{\text{Possible values of } \boldsymbol{\epsilon}} \underbrace{[\int_{\underline{p}_\epsilon}^{r} f(q|C = x_1 + x_2 + \epsilon_1 + \epsilon_2) * q * dq]}_{\text{Average offered price conditional on } \boldsymbol{\epsilon}} d\epsilon_1 d\epsilon_2 \tag{8}$$

Here, as before (in Equation 5), consumers set their reservation price by integrating over the price distribution to get an average price if they search. However, with the added uncertainty over cost, they must also integrate over the distribution of ϵ and calculate that average price for any draw of cost C. With only two firms, we can use the equilibrium density of prices from Subsection 3.1.1 to substitute in for $f(q|C)$ in Equation 8 and solve for r:

$$r = x_1 + x_2 + E[\epsilon_1|p = r] + E[\epsilon_2|p = r] + \frac{s}{A(\mu)} \tag{9}$$

This is very similar to the expression when signals were perfect. In fact, it nests that case, because when signals are perfect, $\boldsymbol{x} = \boldsymbol{c}$ and $\epsilon = 0$. The expectation terms are constants (they do not depend on the actual draw of the ϵ_j terms) which I denote $\hat{\epsilon}_j \equiv E[\epsilon_j|p = r]$. Assumption 2 ensures that the distribution of these ϵ terms does not vary with $\boldsymbol{x}$ which means no moment of the distribution, such as $\hat{\epsilon}_j$, will either. Therefore, conditional on signals $\boldsymbol{x}$, r is fixed, and not a function of the other part of the realized cost ϵ.

This forms an equilibrium because the value of search is decreasing in additional price quotes. Then firms never want to play prices above r because non-shopper consumers search and they lose any shoppers in competition. Consumers never search because they never see a price above r which is the lowest the reservation price could be.

With this reservation price r, we can again find the density of transacted prices, and then integrate over this density to solve for the average transacted price:

$$\bar{p} = x_1 + x_2 + \mu(\epsilon_1 + \epsilon_2) + (1 - \mu)\left(\hat{\epsilon}_1 + \hat{\epsilon}_2 + \frac{s}{A(\mu)}\right) \tag{10}$$

We can see from this expression that the pass-through rate of costs captured by the signal $\boldsymbol{x}$ is 1. However, the part of the cost the consumer is unaware of ($\boldsymbol{\epsilon}$) is only passed through

[7] I write the following expressions for the case of two components of cost ($J = 2$). One could easily rewrite these for the more general case without changing any results.

at a rate of μ.

Proposition 2 defines pass-through for a given pair of consumers' expectations over and realization of the production cost. However, to fully characterize the relationship between changes in cost and changes in equilibrium prices, one needs to define the relationship between costs and consumers' expectations over these costs. Equation 1 breaks the change in cost down into its two components, consumers' expectations x and the residual ϵ. Pass-through is then the change in each of these components multiplied by how price changes with response to these components. Proposition 2 defines the latter relationship and allows us to rewrite Equation 1 as:

$$\rho_j = \frac{\partial \bar{p}(c)}{\partial c_j} = \frac{\partial x_j}{\partial c_j} + \mu \left(1 - \frac{\partial x_j}{\partial c_j}\right) \tag{11}$$

This expression defines pass-through for any model of consumer belief formation about costs. For example, Proposition 1 assumes that consumer information is perfect, or that $c_j = x_j$. With this assumption $\frac{\partial x_j}{\partial c_j} = 1$. Plugging this into Equation 11, one recovers a pass-through rate of 1, which is the conclusion of Proposition 1. As another extreme example, consider the case in which the production cost is always drawn independently and consumers do not observe any information about it. Here consumers' expectations are constant, $x_j = \bar{c}_j$. Therefore, $\frac{\partial x_j}{\partial c_j} = 0$ and $\rho_j = \mu$. In the following section, I explore a model that defines a plausible relationship between changes in cost and changes in consumers' expectations. In Subsection 3.3, consumers update their beliefs rationally in response to noisy signals about an evolving cost. This model defines pass-through as a function of the precision of these signals and nests the above perfect information and no information examples as special cases (resulting in average $\rho \in [\mu, 1]$).

Proposition 2 also allows us to compare pass-through across different components of cost, over which consumers may have different information and beliefs.

Corollary 2.1 *The pass-through rate of two components of cost c_1 and c_2 are equal if and only if $\frac{\partial x_1}{\partial c_1} = \frac{\partial x_2}{\partial c_2}$.*

Pass-through of any given component $c_j = x_j + \epsilon_j$ is then a linear combination of 1 and μ. The pass-through rate of each component is then determined by the ratio of the change in x_j to the change in c_j. As consumers can have different information about different components of cost, this allows for the possibility of different pass-through rates on different components of the marginal cost of production, a novel prediction of this model. In standard models of competition over price, all components of cost are passed through at the same rate[8]. Corollary 2.1 motivates the primary empirical test of this model.

3.3 Dynamic Signals - Information Precision and Pass-through

In many real markets, consumers receive a stream of noisy information about the ever evolving costs facing firms. In this section, I model consumers who continually update their beliefs about firms' costs as they process signals. From this analysis, we get the key prediction of this paper, that pass-through is increasing in the precision of consumers' information about firms' costs.

Consider two infinitely lived firms with short-lived consumers who can observe all signals about marginal costs, but not prices. Consumers see a long series of signals, but only exogenously enter the market to purchase the good (and see prices) in a single period. Costs change and signals arrive in discrete time periods. Assume that the cost evolves according to a Gaussian process. For simplicity of notation, assume there is only one component of cost in each time period c_t.

In order to meet Assumption 1, this cost is transformed using a function with a bounded domain. Put another way, there exists a "cost index," which I denote $c_t^\dagger$, which evolves according to a random walk. Therefore $c_t^\dagger = c_{t-1}^\dagger + \nu_t$, where ν_t is a mean zero Gaussian shock. Consumers know the variance of ν_t, which we denote σ_ν^2. This index is then transformed according to a function, $T : \mathbb{R} \to (\bar{c} - \gamma, \bar{c} + \gamma)$, to get the actual cost c_t. This function is bounded on the interval $(\bar{c} - \gamma, \bar{c} + \gamma)$ where these two values are set such that Assumption 1 is always met. We assume that $T(c^\dagger) = \bar{c} + t(c^\dagger)$. Where $t(0) = 0$, $\lim_{c^\dagger \to -\infty} t(c^\dagger) = -\gamma$, $\lim_{c^\dagger \to \infty} t(c^\dagger) = \gamma$, $t'(c^\dagger) > 0, \forall c^\dagger$, $c^\dagger < 0 \implies t''(c^\dagger) \geq 0$, $c^\dagger > 0 \implies t''(c^\dagger) \leq 0$ and $t(c^\dagger) = -t(-c^\dagger)$.

This gives us a transformation function that effectively truncates the tails of the distribution. This preserves the strict monotonicity of the transformation while bounding the support of the cost distribution. A number of well-known transformation functions fit this assumption including the normal cdf or the logistic function.

In each period, consumers receive signals about this cost index, but never observe the actual cost. Consumers observe a signal y_t and update their beliefs about $c_t^\dagger$ (and c_t). This

[8]Note that this is true even in models where consumers have varying awareness of different components of the price. Agarwal, Chomsisengphet, Mahoney, and Stroebel (2014), for example, provides a simple framework for studying such markets. In such a model, salience of components of *price* can affect the overall level of markups, but changes to any components of *cost* gets passed through at the same rate. Corollary 2.1, therefore, makes empirical predictions that are unique to this model.

signal is the actual index plus noise, or $y_t = c_t^\dagger + \xi_t$, where ξ_t is a mean zero Gaussian shock with known variance σ_ξ^2.

Because the transformation from index to cost (or cost signal) is deterministic and the function is invertible, inference over costs is equivalent to inference over the index, with the appropriate transformation. Given this setup, the consumers' beliefs over $c_t^\dagger$ can be represented by a Kalman filter and can be fully summarized at any time t by their estimate of $c_{t-1}^\dagger$, $\hat{c}_{t-1}^\dagger$, and the estimated variance of this estimate, $\hat{P}_{t-1}$. These two estimates completely capture all information up to time t. Assume that at time t, $c_{t-1}^\dagger = 0$ and thus $c_{t-1} = \bar{c}$.

Proposition 3 *When costs and signals about costs are generated according to the above process, a reservation price equilibrium exists and average pass-through (evaluated at $c_{t-1} = \bar{c}$) is between μ and 1. It is decreasing in the variance of the signal σ_ξ^2 and increasing in the number of shoppers in the market μ.*

To start, I show how consumers' beliefs evolve. At time t, the individual expects the cost index to be $\hat{c}_{t|t-1}^\dagger = \hat{c}_{t-1}$ and the variance of this is $\hat{P}_{t|t-1} = \hat{P}_{t-1} + \sigma_\nu^2$. Upon seeing a signal, the individual updates their beliefs using the Kalman gain term:

$$K_t = \frac{\hat{P}_{t-1} + \sigma_\nu^2}{\hat{P}_{t-1} + \sigma_\nu^2 + \sigma_\xi^2} < 1. \tag{12}$$

This term captures the degree to which consumers update their beliefs upon receiving new information. After receiving the signal their expected cost index is $\hat{c}_t^\dagger = (1-K_t)\hat{c}_{t-1} + K_t y_t$. We can use this gain term to solve for the steady state variance of the consumers' estimate of the production cost index, which I denote $\hat{P}$. The variance of consumers' beliefs is increasing in the variance of the signal (σ_ξ^2) and the variance of the shocks to the underlying cost index (σ_ν^2). It is worth noting that if *either* consumers perfectly observe costs ($\sigma_\xi^2 = 0$) or costs evolve according to a deterministic process ($\sigma_\nu^2 = 0$), than this variance is zero, and the model simplifies to the perfect information Stahl (1989) case. We can also plug this steady state variance into Equation 12 to get the steady state Kalman gain term $\hat{K}$ which is less than one and is decreasing in σ_ξ^2. This defines the signal structure and updating process of the cost index.

To present the intuition behind this proposition, I first show that the statements about pass-through hold when costs are unbounded and the cost index is equal to the actual cost. This violates Assumption 1 and a reservation price equilibrium does not exist when costs are distributed accordingly. However, the lack of transformation makes the algebra easier to follow, and the same logic applies in the bounded case. A formal proof that the equilibrium exists in the bounded case, and that the same statements about pass-through hold, is presented in Online Appendix A. In this exposition, we are assuming that the actual marginal cost of production equals this cost index. Assume that consumers have observed these signals for a sufficient period of time and thus their beliefs over costs are normally distributed with mean $\hat{c}_{t-1}$ and variance $\hat{P}$. We assume that consumers' mean beliefs are correct at the time of analysis and thus $c_{t-1} = \hat{c}_{t-1} = \bar{c}$.

If a reservation price existed for these beliefs (as it does when the cost index is transformed as above), then the average price for any realization of cost (c) and noise (ξ) is defined by Equation 10 which can be rearranged slightly as follows:

$$\bar{p}(c, \xi) = x(c, \xi) + \mu(c - x(c, \xi)) + (1 - \mu)\left(\hat{\epsilon}(c, \xi) + \frac{s}{A(\mu)}\right) \tag{13}$$

Here I have rewritten ϵ as $c - x$ (the actual realization of the cost less the consumers' expectation of that cost). I have also expressed x, the mean of consumers' beliefs over costs, and $\hat{\epsilon}$, the mean of consumers beliefs over ϵ, conditional on seeing a price of r, as functions of c and ξ. Both of these values are actually just functions of the signal that the consumer received y. However y can be decomposed into the actual cost and the noise term ξ. This decomposition will be useful when we integrate over possible signals. The average price given cost c is then the average price across all possible realizations of ξ:

$$\bar{p}(c) = \int_{-\infty}^{\infty} [x(c, \xi) + \mu(c - x(c, \xi)) + (1 - \mu)\left(\hat{\epsilon}(c, \xi) + \frac{s}{A(\mu)}\right)]g(\xi)d\xi \tag{14}$$

Where $g(\xi)$ is normal with mean 0, variance σ_ξ^2. Average pass-through is then the derivative of this average price with respect to cost:

$$\rho(c) = \int_{-\infty}^{\infty} [x_1'(c, \xi) + \mu(1 - x_1'(c, \xi)) + (1 - \mu)\hat{\epsilon}_1'(c, \xi)]g(\xi)d\xi = \mu + (1 - \mu)\int_{-\infty}^{\infty} [x_1'(c, \xi) + \hat{\epsilon}_1'(c, \xi)]g(\xi)d\xi \tag{15}$$

When beliefs are unbounded, Assumption 2 is met and $\hat{\epsilon}_1'(c, \xi) = 0$ for all c and ξ.[9] Thus,

[9] This will not be the case when the distribution of c is bounded. However, as is shown in the appendix, this value being non-zero does not change either of the conclusions made in Proposition 3

for this exposition, the expression for average pass-through simplifies to:

$$\rho(c) = \mu + (1-\mu)\int_{-\infty}^{\infty} x_1'(c,\xi)g(\xi)d\xi \tag{16}$$

To understand this expression, it will be useful to characterize the consumers' beliefs over cost given an actual cost draw and the realization of the noise in their signal, $x(c,\xi)$. First, we know that given our assumed prior $\hat{c}_{t-1}$, a new cost c, and a noise shock ξ, consumers' will believe that the cost $c^\dagger$ is distributed normally with mean

$$x(c,\xi) = \hat{K} * (c+\xi) + (1-\hat{K})\hat{c}_{t-1} \tag{17}$$

and variance $\hat{P}$. The derivative of this expression with respect to c is then simply $\hat{K}$. Plugging this into Equation 16, we get that average pass-through is:

$$\rho(c) = \mu + (1-\mu)\hat{K} \tag{18}$$

This expression is clearly increasing in $\hat{K}$. As $\hat{K}$ is decreasing in the variance of the signals that consumers receive, σ_ξ^2, pass-through is decreasing in this variance as well. Furthermore, as $0 \le \hat{K} \le 1$, pass-through is between μ and 1 and is increasing in μ. While the transformation that bounds costs makes the expression for pass-through more complicated, it does not change either of these results. Therefore, average pass-through is decreasing in the variance of consumers' signals and moves towards one as the precision of their information increases.

This model also gives us predictions about the probability of observing any given pass-through rate. In the previous exercises, I discuss average pass-through, the derivative of average transacted price with respect to marginal cost, calculated by integrating over all possible realizations of the signal that consumers receive. Another object of interest might be the probability of observing a certain pass-through rate. I define an observed pass-through rate as the change in average transacted price following a change in marginal cost, divided by that cost. As an example, given recent discussions in the media, one might be interested in the probability of observing a pass-through rate greater than one, or over-shifting.

Corollary 3.1 *When consumers perfectly observe costs or have no information about costs, $\sigma_\xi^2 \in \{0,\infty\}$, the probability of observing pass-through greater than one is 0 (evaluated at $c_{t-1} = \bar{c}$). When consumers have imperfect information about costs, $\sigma_\xi^2 \in (0,\infty)$, the* *probability of an observed pass-through rate greater than one is strictly greater than 0.*

Again, I show that this statement holds when costs and beliefs about costs are unbounded. That this continues to hold when costs are bounded and Assumption 1 is met is left to the proof in the appendix. The intuition behind this corollary is captured by this example despite the non-existence of a reservation price equilibrium.

That the probability of over-shifting equals zero when $\sigma_\xi^2 = 0$ follows directly from Proposition 1. When consumers have perfect information about the cost, as they do when they observe noiseless signals, the pass-through rate equals one with certainty.

For $\sigma_\xi^2 \neq 0$, we are again assuming that $\hat{c}_{t-1} = c_{t-1} = \bar{c}$. Assume that cost has increased by $\delta > 0$. Then the change in price can be written as a function of ξ (and σ_ξ^2). When costs are unbounded, Assumption 2 is met and $\Delta\hat{\epsilon} = 0$. Given the simple belief updating structure, the change in consumers' expectations over costs reduces to $\hat{K}y$ or $\hat{K}(\delta + \xi)$, resulting in the following price change:

$$\Delta\bar{p}(\xi, \sigma_\xi^2) = \mu\delta + (1-\mu)\hat{K}(\delta + \xi) \tag{19}$$

This expression is clearly strictly monotonic in ξ. Therefore if there exists a value of ξ, ξ^* for which $\Delta\bar{p}(\xi^*, \sigma_\xi^2) = \delta$, it must be that $\Delta\bar{p}(\xi, \sigma_\xi^2) > \delta$ for all $\xi > \xi^*$ and $\Delta\bar{p}(\xi, \sigma_\xi^2) < \delta$ for all $\xi < \xi^*$. In other words there exists a cutoff ξ^* such that for any noise term greater this cutoff, pass-through will be greater than 1. We can set Equation 19 equal to δ to solve for this cutoff:

$$\xi^* = \frac{1-\hat{K}}{\hat{K}}\delta = \frac{\delta\sigma_\xi^2}{\hat{P} + \sigma_\nu^2} \tag{20}$$

The noise term ξ is distributed normally with mean zero and variance σ_ξ^2, so the probability of a realization of $\xi > \xi^*$ is $1 - \Phi(\frac{\xi^*}{\sigma_\xi}) = 1 - \Phi(\frac{\delta\sigma_\xi}{\hat{P}+\sigma_\nu^2})$, where Φ is the standard normal distribution.[10]

For $\sigma_\xi^2 \in (0, \infty)$, the value of $\frac{\delta\sigma_\xi}{\hat{P}+\sigma_\nu^2}$ is positive and finite. This means that $F(\frac{\delta\sigma_\xi}{\hat{P}+\sigma_\nu^2}) < 1$, and thus the probability of over-shifting is strictly greater than zero.

When $\sigma_\xi^2 = \infty$, the term $\frac{\delta\sigma_\xi}{\hat{P}+\sigma_\nu^2}$ is not finite. To see this, note that the numerator increases linearly in σ_ξ. The denominator, on the other hand, contains the term $\hat{P}$ which increases in σ_ξ at a rate less than linear. This means that the probability of observing a signal $\frac{\xi^*}{\sigma_\xi}$ greater than this value is zero. The probability of over-shifting is then zero when $\sigma_\xi^2 = \infty$.

[10]One cannot make this normalization when $\sigma_\xi^2 = 0$. In that case, the result follows directly from Proposition 1

This corollary shows that when consumers have either perfect information about costs or no information about costs, observed pass-through cannot be greater than one. However, when consumers have any intermediate level of imperfect information about costs, overshifting occurs with positive probability.

3.4 Firm Revelation of Cost - Asymmetric Pass-through

A natural extension of this model is to allow firms to communicate the costs that they are facing to potential consumers. Consider an extension to the model in which firms can (at a cost d) reveal their marginal cost of production to consumers who have requested a price quote. This model formalizes the discussion of firm revelation of cost in Section 4 of Janssen, Pichler, and Weidenholzer (2011), analyzes pass-through in this setting, and explains the empirical phenomenon of asymmetric pass-through.

For there to exist a non-trivial equilibrium in which firms reveal their costs for some realizations and not others, we must place bounds on the cost of revelation.

Assumption 3 *Firms can pay a cost d to reveal their marginal costs. This revelation is verifiable and d lies on the interval $(0, \frac{1-\mu}{2}(\bar{\epsilon} - E[\epsilon|p = r])]$.*

Then, when this assumption is met, the following Proposition holds.

Proposition 4 *When Assumptions 1, 2, and 3 are met, there exists an equilibrium such that for a given signal x, there is a value of actual cost c^* above which all firms reveal their costs to consumers and below which they do not. This results in an average pass-through rate equal to one when costs are above c^*, and an average pass-through rate lower than one when costs are below c^*.*

This version of the model has a separating equilibrium. When the difference between firms' costs and consumers expectations (ϵ) is above a certain value ϵ^*, all firms will choose to pay the disclosure cost d and reveal their costs to consumers. Consumers then know firms' actual costs with certainty. When costs are such that $\epsilon \leq \epsilon^*$, firms will not reveal their costs, leading consumers to update their beliefs about firms' costs, knowing that $\epsilon \leq \epsilon^*$.

If the firms reveal their costs, the consumer has full information and $\epsilon = 0$. Then, by the results from Equation 7, the pass-through rate is 1. Upon not seeing a revelation of costs, consumers will update their beliefs about c, changing the x and ϵ that go into the calculation of average price in Equation 10. Consumers' expectations about cost become $x + E[\epsilon|\epsilon \leq \epsilon^*]$. This expectation term is negative because the unconditional expectation is zero. This then gives us 2 regions with different properties of pass-through. As stated above if $\epsilon \geq \epsilon^*$, $\rho = 1$. If $\epsilon < \epsilon^*$, then the average pass-through rate must be less than one, because the average signal is negative.

This generates an asymmetry of pass-through around the cost $c^* \equiv x + \epsilon^*$. More generally, this extension to the model implies that pass-through of higher than expected costs is greater than pass-through of lower than expected costs. If costs are dynamic as in the previous section, this is consistent with the well documented empirical fact that "prices rise faster than they fall (Peltzman (2000))."

4 Data and Setting

4.1 Setting and Institutional Details

I study the role of consumer beliefs in determining pass-through in the US residential mortgage market. This market is ideal for testing the model's prediction that different components of marginal cost can have different pass-through rates if consumers have different information about each component. Critically, the mortgage market features both search frictions and multiple observable costs of which consumers have varied awareness. I use the confidential Home Mortgage Disclosure Act data to show that widely known costs are passed through at a higher rate than more obscure ones.

The majority of mortgages in this market are 30-year, fixed-rate loans that are guaranteed by the US government, securitized into a MBS and sold to investors. This securitization and sale is the primary source of funding for lending and is the largest determinant of the marginal cost of lending to a borrower. When a lender originates a loan for borrower i at time t, securitizes, and sells it, they receive the following upfront cash flows[11]:

$$\omega_{it} = \text{TBA Price}_t + \text{Spec Pool Payup}_{it} - \text{LLPA}_{it} - \text{points}_i - \text{buydown}_i - \text{unobservable costs}_{ijt} - 100$$

This cash flow, ω_{it} is quoted in dollars per hundred dollars lent to the borrower. I will focus primarily on three of these components: TBA price, specified pool payup and LLPA[12]. Im-

[11] Adapted from Fuster, Goodman, Lucca, Madar, Molloy, and Willen (2013)

[12] Points refers to an upfront payment made by the borrower in exchange for a lower interest rate. In the empirical analysis, I will define notions of price that are inclusive of points taken. Buydown refers to the firms' ability to make an upfront payment to the government sponsored entities to lower the recurring guarantee fee they pay. I assume that such a transaction has no effect on the discounted present value of the cost to the lender, and can therefore be ignored. Unobserved costs include the many costly actions that

portantly, all three of these components are observable to the econometrician and common across all lenders. I will address each of them in turn.

The majority of mortgages that are securitized are sold into an MBS market known as the to-be-announced (TBA) market[13]. The TBA market is an incredibly liquid, futures-like market in which parties exchange standardized contracts outlining the future exchange of a MBS for a set price. These contracts do not specify the exact mortgages that must be delivered but rather broad characteristics such as the governmental agency guaranteeing the loan and the coupon rate of the bond. The flexible nature of these contracts and the liquidity provided by this market make TBA trades the primary funding source for lending.

This market has very transparent pricing and the yield on these securities closely tracks most market interest rates such as the ten-year US Treasury yield. Figure 1 plots the yield on these securities alongside the yield on the ten-year US Treasury note and shows that the two move in tandem. The correlation coefficient on the two yields is 0.95 during the period studied in this paper. While borrowers may not understand the exact funding process for their loans, they are likely to be well aware of the prevailing rates and have a strong sense of this main component of firms' cost of lending.

Figure 1: United States 10-year Treasury Note Yields and Bank of America Fannie Mae 30 Year TBA Current Coupon Yield plotted over the studied period.

---

lenders must take in order to complete the lending and securitizaton process. This includes the time of the loan officer, legal fees associated with securitization, marketing costs, and many other costs. I assume that, conditional on a number of observable covariates, these are orthogonal to the costs being studied. Lastly, the 100 in the equation refers to the face value of the loan, paid to the borrower by the lender.

[13]TBA trades make up around 90% of daily trades in the agency MBS market (Fuster, Lucca, and Vickery (2023)).

The specified pool payup is an additional payment that lenders receive for loans that have certain characteristics. MBS are fixed rate securities, so their value decreases when market interest rates increase. However, they have the unique property that they are negatively convex, meaning the second derivative of their price with respect to rates is negative. When rates fall, mortgage borrowers prepay and refinance their loans, returning money to investors at par. Similarly, when rates rise, borrowers are reluctant to prepay, lengthening the expected duration of the bond. This causes the value of an MBS to decrease faster and increase slower with respect to market rates than a similar, non-convex bond. To compensate investors for this *prepayment risk*, MBS trade at a lower price and thus a higher yield than these similar bonds. There are certain characteristics of mortgages that can reduce this prepayment risk and thus reduce the convexity of the bond, increasing its price. When an MBS contains many mortgages that all feature one of these properties, it is typically traded as a specified pool and is traded at a premium to TBA. The most common example of such a feature is *low loan balance*. Borrowers with a low loan balance have a lower incentive to refinance because the rate incentive needs to be larger to offset the fixed costs of refinancing. Therefore investors prefer these sorts of borrowers.

Market conventions have settled on trading loans with low loan balances by "buckets" of loan sizes. For example, Low Loan Balance or LLB bonds exclusively contain mortgages with a loan balance below $85,000. Medium Loan Balance or MLB bonds contain loans with balances between $85,000 and $110,000. These different categories of MBS are commonly traded up to a category containing loans between $175,000 and $200,000. Naturally, a LLB bond has lower prepayment risk than a MLB bond, and thus trades at a higher price. A lender therefore gets greater cash flow at origination, a higher "payup" to TBA, and thus has a lower marginal cost of lending, for loans that can be placed in a better specified pool. This creates a discontinuity in the marginal costs faced by lenders with respect to loan size. An $84,000 loan is very similar to an $86,000 loan in every sense, besides the fact that the $84,000 loan can be placed in an LLB bond while the $86,000 cannot, meaning the marginal cost of lending to the $84,000 borrower is, all else being equal, lower. Borrowers are unlikely to be aware of the market for specified pools. Furthermore, the sign of the effect of increasing loan size and moving from one specified pool bucket to another is the opposite of what intuition would lead borrowers to believe. Generally, increasing loan size spreads the fixed costs of lending over a larger loan, lowering the cost of lending. However, near these loan-size cutoffs, an increase in loan size causes a drop-off in pay-ups, resulting in a discrete increase in the lenders' cost, unbeknownst to the borrower.

Loan level price adjustments (LLPAs) are payments made by the lender to Fannie Mae and Freddie Mac to guarantee the credit risk of the loan upon securitization. The GSEs require greater upfront payment for riskier borrowers, but they make this assessment based on discrete buckets of credit scores and loan-to-value (LTV) ratios, creating a "grid" of LLPA values[14]. This, again, creates a discontinuity in the marginal cost of lending to borrowers in the credit score and LTV dimensions. For example, consider two identical borrowers taking out loans with an LTV of 80%, one of which has a credit score of 699, the other 701. The true credit risk of these borrowers is more or less indistinguishable but a lender to the 699 borrower needs to pay an LLPA of 1.75% while a lender to the 701 borrower needs to pay an LLPA of 1.25%. Therefore, lenders face a lower marginal cost of lending when borrowers are below one of these thresholds. Borrowers may be unfamiliar with the exact details of how LLPA is determined, but they are likely to be aware that lower credit scores and lower down payments (higher LTV ratios) result in higher costs to lenders.

4.2 Data

I use the confidential version of the Home Mortgage Disclosure Act (HMDA) database from 2018-2020. These data contain almost all residential loans made in the US during this time period and contain critical fields such as the loan application date, loan amount, interest rate, other payments and fees, credit score, loan-to-value ratio, debt-to-income ratio, lender, home location (3 digit zip code), and other borrower demographics. I exclude the pre-2018-2020 period because interest rates were not previously recorded in HMDA. To study a homogenous product, I also restrict attention to purchase loans that are conventional, conforming, 30-year, fixed-rate, single-family, owner-occupied, and standard-amortization[15].

The confidential version of the data provides two advantages relative to the publicly available data set. First, it provides credit scores and exact loan amounts, which allow for exact identification of LLPAs and specified pool payups. Second, it provides exact loan application dates rather than loan application months or quarters, which allows for the use of daily TBA prices and specified pool payups. This reduces measurement error in marginal costs which is critical for identifying pass-through rates without attenuation bias. I link the

[14]For current LLPAs see grid.

[15]This segment of the mortgage market was chosen because it is the largest portion of the market and is closest to what market participants would consider "standard mortgages." I have not performed any of the following analysis on any other subset of the market.

confidential HMDA data to eMBS securities level data, which records the specific MBS pools that loans have been placed in. This allows me to attribute specified pool payups to loans that were *actually* placed in specified pools rather than just those that are *eligible* to be placed in specified pools. For TBA and specified pool payup pricing, I use daily historical TBA prices and specified pool payups from Citigroup's Citi Velocity platform[16]. Finally, for LLPAs, I use Fannie Mae's LLPA matrix.

4.3 Converting Interest Rates to Prices

To estimate the pass-through rate of marginal cost to equilibrium prices, the two must be measured in the same units. In most of the literature, this is straightforward, both cost and price can be expressed in dollars. However, in the mortgage market, the costs that we are studying are each upfront payments to or from the lender. The most prominent element of the price, on the other hand, is the interest rate. Therefore, to estimate pass-through, I must convert the payments made by the borrower into a measure of present value. Below, I suggest one such method of doing so, using present value calculations.

My primary method of calculating the price paid by the borrower is to calculate a present value of the loan to the borrower. The borrower receives the face value of the loan upfront, less fees and points paid, but then pays a monthly payment based on the interest rate of the loan for the life of the loan. If we assume a discount rate and a prepayment schedule, we could calculate the present value of all of these cash flows. Here I use the industry standard PSA prepayment model which assumes borrowers do not prepay at all in the first month, ramp up their prepayment over the next 30 months linearly, and then prepay at a constant rate for the remaining months. This is of course a simplification of borrowers' prepayment behavior, but it is used as a benchmark by market participants and approximates actual payments. In my main specification, I assume an annual discount rate of 10% and assume prepayment of 250 PSA. I then calculate the cash flows in each month under this prepayment assumption and discount them to get a present value. The benefit of using this method is that it reflects how economists think of the total benefit or loss received by a borrower from a change in the interest rate. The main downside is that it requires us to make assumptions about the discount rate and prepayment behavior.

In the following section, I estimate pass-through using this measure of price. In Online

[16]This platform did not report payups for the more infrequently traded specified pools with loan size less than $200,000. To calculate payoffs for loans that were placed in these pools, I extrapolate from the payups of $150,000 and $175,000 cutoff pools.

Appendix B, I suggest two additional methods of calculating price: using prices facing borrowers and using prices facing lenders. I repeat the below analysis using these notions of price (and find qualitatively similar results).

5 Results

5.1 Empirical Strategy

The main goal of this empirical exercise is to measure and compare the pass-through rates of the three components of the marginal cost of lending. To do this, I first run what Muehlegger and Sweeney (2021) call the "canonical pass-through regression," but modified to include multiple components of marginal cost. Specifically, I regress (in levels):

$$\text{Price}_{ijlt} = \rho_1 \text{TBA}_t + \rho_2 \text{Payup}_{it} + \rho_3 \text{LLPA}_i + \beta X_i + \nu_l + \epsilon_{ijlt} \tag{21}$$

Where ρ_1, ρ_2, and ρ_3 are the pass-through rates being studied. The dependent variable, Price_{ijlt}, is the transaction price between borrower i and lender j in county l on day t. To estimate pass-through of cost to price, this price must be measured in the same units as TBA_t, Payup_{it}, and LLPA_i, dollars paid upfront per hundred dollars of loan. This presents a challenge, as the "price" facing borrowers is typically quoted as an interest rate. To address this issue, I use the aforementioned methods of converting all of the payments made by borrowers into one upfront price and demonstrate that each of these leads to the same conclusion: pass-through is not uniform across these components of cost.

The vector X_{it} contains individual level characteristics, most importantly those that are used to determine Payup_{it} (loan size) and $LLPA_{it}$ (credit score and LTV). Location fixed effects are denoted by ν_l. An unobservable individual level term, ϵ_{ijlt}, allows for correlation within firm, location and time.

This specification allows for the identification of the three pass-through rates under the following assumptions. Parameter ρ_1 is identified if the TBA price is statistically independent from ϵ_{ijlt}. This is a common assumption in the literature (Fuster, Lo, and Willen (2017) and Scharfstein and Sunderam (2016)) and it is plausible given the massive size of the TBA market and its interconnectedness with other global rates markets.

The identification of ρ_2 and ρ_3 is similar to that of a regression discontinuity design (RDD). Using the terminology of Cattaneo and Titiunik (2022), I present identification arguments under a continuity framework and a local randomization framework. In the continuity framework, parameter ρ_2 is identified if the relationship between the unobservable costs of lending and borrowers' loan sizes is continuous around the specified pool cutoffs and loan size is continuously distributed. Parameter ρ_3 is identified under a similar assumption. Unobservable costs must be continuous in credit score and LTV ratio around the cutoffs for different LLPA buckets.

In the local randomization framework, ρ_2 is identified if, for a given borrower wanting to take out a loan for an amount near a cutoff, whether their loan is above or below that cutoff is random. Likewise ρ_3 is identified if, near the LLPA cutoffs, whether a credit score or LTV ratio is above or below the cutoff is random. The identification of these parameters rely on the assumption that the running variables are not strategically chosen by borrowers which I address in more detail below.

Ideally, these ρ_2 and ρ_3 are identified by comparing loans made to individuals directly above and below one of the cutoffs. Local randomization of these the three running variables (loan size, credit score, LTV) is plausible. Secondly, the relationship between these variables and the unobservable elements of price should be continuous at each cutoff. Intuitively, individuals with a credit score of 699 should be similar to individuals with a credit score of 701 in all aspects besides the LLPA they pay. Likewise individuals who take out a \$109,000 loan should be similar to those who take out a \$111,000 loan. Borrowers may make a strategic decision over the down payment they make and thus the LTV ratio of their home. I therefore estimate ρ_2 and ρ_3 both across LTV ratios and holding it fixed at 80, the most common choice made by borrowers.

I use two different approaches to estimation. The first is an RDD with multiple cutoffs. Here, I estimate Equation 21 using only loans that are near a cutoff. In my first specification using this approach, I focus on loans with an LTV of 80% because there is a large mass of borrowers with this LTV ratio as this is the cutoff above which borrowers need to take out mortgage insurance. This has the added benefit of making the local randomization of loan size assumption more plausible. If borrowers are committed to an 80% LTV loan, then their loan size is simply 80% of the homes' purchase price which should be locally random. Credit scores are determined by an external agency and are difficult to game above and below a certain cutoff. I allow for different slopes of β with respect to loan size and credit score around the different cutoffs but estimate only single values of ρ_2 and ρ_3[17]. I

[17] This is accomplished using indicator variables specifying which loan size bucket the loan is in and interaction terms between these indicators and the loan size variable. These terms are added to the matrix X_i

then estimate the same equation on the broader set of loans with an LTV ratio less than 80 that are near one of the LLPA cutoffs. Here I allow for different slopes of β with respect to loan size, LTV, and credit score around the different cutoffs.

The second approach is to jointly estimate these parameters using the entire dataset. Here, it is important that we allow for sufficient flexibility in the relationship between the running variables and price. I therefore estimate a slightly different equation:

$$\text{Price}_{ijlt} = \rho_1 \text{TBA}_t + \rho_2 \text{Payup}_{it} + \rho_3 \text{LLPA}_i + f(\text{Loan Size}_i) + g(\text{FICO}_i) + h(\text{LTV}_i) + \beta X_i + \nu_l + \epsilon_{ijlt} \tag{22}$$

Here I include third order polynomial terms to non-parametrically estimate the continuous relationship between the three running variables and price. Then any change in price caused by the discontinuous jump in payups or LLPA at the cutoffs will be captured by ρ_2 and ρ_3. I estimate this equation on the larger, full sample of loans.

5.2 Hypothesis

Connecting these regression models to the predictions made by the theoretical model of pricing outlined above, the theory predicts that the ρ terms will be functions of the precision of information borrowers have about each component of cost. Assuming consumers have less information about some components (such as specified pool payups) than others, we should expect $\rho_1 \neq \rho_2 \neq \rho_3$. Specifically, if one believes that consumers will have more precise information about rates and LLPAs than specified pool pay-ups, then the model predicts that $\rho_1, \rho_3 > \rho_2$. This is in contrast to other models of equilibrium pricing which predict $\rho_1 = \rho_2 = \rho_3 = \rho$.

5.3 Graphical Results

Before jumping into the results of the regressions above, I visually inspect the relationship between the pricing of two of the components of marginal cost, TBA prices and specified pool payups. Below, I present a series of bin-scatter plots. In each of these, I use the present value of the loan to the borrower (assuming 250 PSA prepayment and a 10% discount rate) as the measure of transacted price[18]. This present value is decreasing in the interest rate

in Equation 21. A similar approach is used with credit scores and/or LTV ratios when estimating pass-through of LLPA.

[18]The use of a single, constant prepayment rate for all borrowers is a simplification of how borrowers view the total price of the loan. Borrowers vary in their propensity for prepayment; however, what is important to

paid by the borrower and should be increasing in TBA prices and specified pool payups. I then residualize this present value by regressing it against controls for county and the LLPA grid. Finally, I plot this residualized present value against either TBA prices or payups. Figure 2 studies the pass-through of TBA prices while Figures 3 and 4 study the pass-through of specified pool payups.

Figure 2 plots residualized transacted present value against the 3.5% coupon TBA price on the day the loan was originated. Present value is residualized against county, loan amount, and LLPA bin fixed effects. One can see that these transacted values are closely tied to the TBA prices. Furthermore, the slope of this line is 1.03 which is the estimated pass-through of TBA prices. Note that throughout these results, the overall levels of pass-through depend on the transformation used to convert interest rates to upfront payments. This study is not concerned with the overall level of pass-through per se, but rather the *relative* levels of pass-through of the different components. This estimated pass-through is similar in magnitude to the nearly complete pass-through found in Fuster, Lo, and Willen (2017) of 0.92. This slope is therefore useful primarily as a benchmark to compare to the pass-through of specified pool payups.

---

the regression-discontinuity-like empirical design, is that any relationship between observable characteristics, such as loan size, and prepayment behavior is continuous. Borrowers with $109,000 loans should not have significantly different prepayment behavior than borrowers with $111,000 loans, for example. Furthermore, the use of a constant prepayment rate assumption should be viewed as a conservative estimate of the test of differential pass-through. If borrowers with smaller loan sizes prepay more slowly than borrowers with larger loan sizes, than my estimates for the total prices paid by small loan borrowers are biased downward. Lenders to these small loan borrowers receive larger payups, therefore the estimates for pass-through of payups would be biased upwards, making it more difficult to reject the null hypothesis that the pass-though of TBA prices and payups are equal.

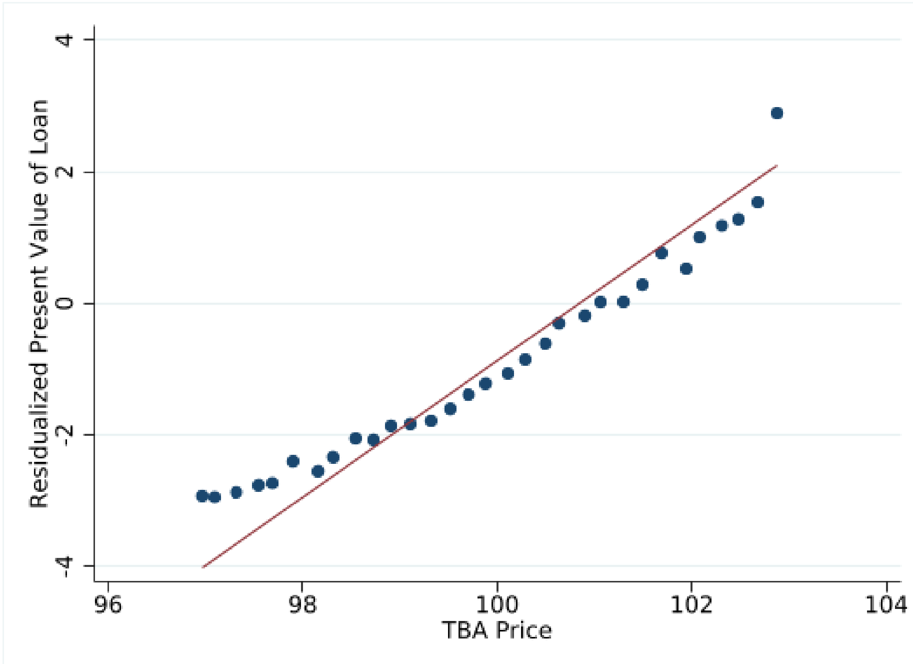
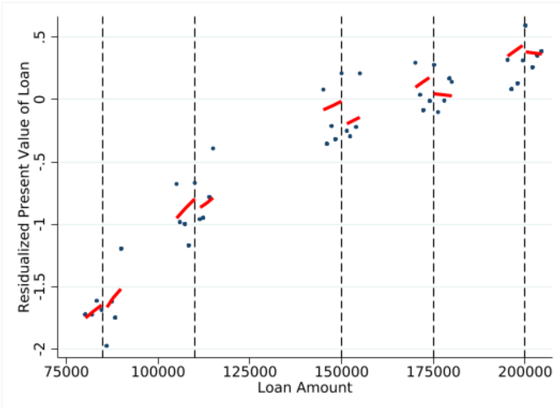
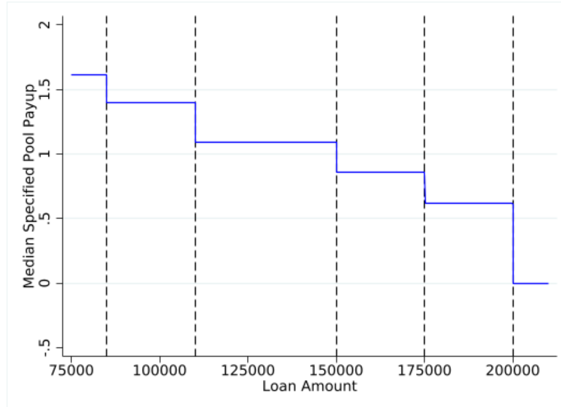


Figure 2: Transacted borrower present values less county, loan amount, and LLPA bin fixed effects plotted against 3.5% coupon TBA price on the day of loan origination. Solid line plots the line of best fit, representing the pass-through of TBA price to borrower prices.

Next, Figure 3a plots residualized borrower present values against the borrower's loan size. Present values are residualized against county, date, and LLPA bin fixed effects. As discussed above, loan size determines the specified pool that loans can be placed in and thus determines the specified pool payup that the lender receives. As loan size increases, the loan becomes ineligible for higher payup pools and thus the payup the lender receives decreases. Figure 3b shows how the average payup earned by lenders varies across loan sizes. The payup difference across these cutoffs is on average about \$0.25 per \$100 of face value. While Figure 3a shows that there are clear downward jumps in borrower present value at each of the loan size cutoffs, the decreases in present value to the consumer are small relative to the loss in payup to the lender. The sizes of these jumps in borrower present value relative to the changes in payups provide a visual representation of the regression discontinuity approach to estimating the pass-through of specified pool payups.



(a) Transacted borrower present values less county, application date, and LLPA bin fixed effects. Solid lines plot the lines of best fit within each LLPA bucket. Vertical lines show specified pool loan size cutoffs.



(b) Average specified pool payups received by lenders plotted against loan size. Vertical lines show specified pool loan size cutoffs.

Figure 3: Regression discontinuity plots of prices and costs against loan size.

To compare the pass-through of specified pool payups directly to that of TBA prices, Figure 4 plots residualized borrower present values against the specified pool payup for that loan. Present values are residualized against county and LLPA bin fixed effects as well as linear effects of TBA prices and the size of the loan (where the effect is allowed to vary by specified pool bucket as in an RD regression). This figure only uses loans for borrowers with loan sizes close to the specified pool cutoffs to isolate the effect of specified pool payups on price. The slope of the line of best fit is 0.23, which is the estimate for pass-through of payups to borrower present value. The dashed blue line in this figure represents what the relationship between payups and borrower present value would be if the pass-through of specified pool payups was the same as that of TBA prices. Comparing the line of best fit to this dashed line, it is clear that the pass-through rate of specified pool payups is far lower than that of TBA prices.

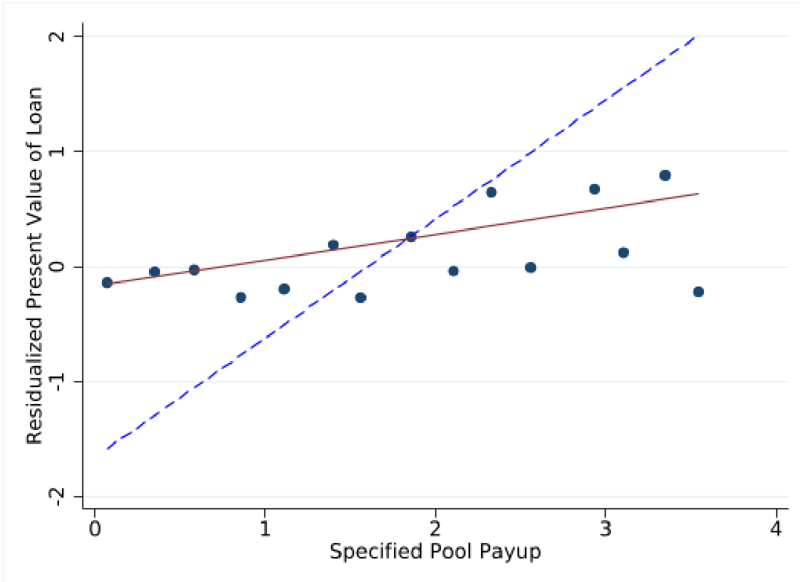


Figure 4: Transacted borrower present values less county and loan amount fixed effects with linear effects of TBA price and linear effects of loan size that vary across specified pool buckets plotted against specified pool payups. Plot only contains data points within $5,000 of a specified pool cutoff. Solid line plots the line of best fit, representing pass-through of specified pool payups to borrower prices. Dashed blue line plots the hypothetical relationship between specified pool payups and prices if payups had the same pass-through rate as TBA prices.

The identification strategy for LLPA pass-through is very similar to that of specified pool payups. Instead of using discrete cutoffs in loan sizes used by the specified pool market, we use the discrete buckets of credit scores and LTV ratios used to determine the LLPAs set by the GSEs.

Based on the above figures, it appears that the pass-through of specified pool payups is far lower than that of TBA prices. This difference in the relative pass-through rates is consistent with the theoretical model above, but not with other models of equilibrium pricing. Below, I confirm this using the regression discontinuity regressions discussed above.

5.4 Regression Results

Table 1 estimates pass-through of the three different components of cost using borrower present value as calculated above as the notion of price. In each column, I report the point estimate for pass-through of these components under a different specification. Below the point estimate is the 95% confidence interval for that pass-through rate. Finally, at the bottom of each column, I report the p-values for a F-test that compares the coefficients (pass-through rates) of TBA price vs specified pool payups or LLPA. When this value is below a critical value, we can reject the null hypothesis that the components have the same pass-through rate. This null hypothesis is the hypothesis under a standard price competition model.

Column 1 of Table 1, reports the results of estimating Equation 21 without the LLPA terms, isolating the estimation of ρ_1 and ρ_2. The equation is estimated using only loans that are within $5000 of a specified pool cutoff. I control for loan size linearly but with the relationship between loan size and cost allowed to vary across loan size buckets. I then control for county and LLPA bucket fixed effects. Under this specification, I receive point estimates of $\rho_1 = 0.85$ (TBA price pass-through) and $\rho_2 = 0.37$ (Specified pool payup pass-through). These estimates are significantly different at a 1% significance level. We can therefore be confident that specified pool payup pass-through is significantly lower than that of TBA prices.

Column 2 reports the results of estimating Equation 21 without the payup terms, isolating the estimation of ρ_1 and ρ_3. Here, the equation is estimated only on loans with an LTV ratio of 80 that have a FICO score within 5 of the LLPA cutoffs. This is done to estimate the pass-through of LLPA using only the variation in the borrowers credit score. I estimate this separately as credit score is more likely to meet the assumption of being locally random near the cutoffs than LTV ratio. In this regression, I control for credit score linearly but allow the relationship to vary across LLPA buckets. I include county and loan size bucket fixed effects. Here, I receive point estimates of $\rho_1 = 0.97$ and $\rho_3 = 0.87$ (LLPA pass-through) although the difference between these rates is not significant. We therefore cannot reject the null hypothesis that TBA prices and LLPAs are passed through at the same rate.

Column 3 estimates the same equation as column 2, but uses the larger sample of loans that are either within 5 of a credit score cutoff or within 2 of an LTV ratio cutoff. I use the same fixed effects but introduce additional linear controls for LTV ratio that can vary across LLPA buckets. This gives us very similar point estimates of $\rho_1 = 0.99$ and $\rho_3 = 0.96$. Again, we cannot reject the null hypothesis that TBA prices and LLPAs are passed through at the same rate.

Column 4 estimates the full Equation 21, jointly estimating ρ_1, ρ_2, and ρ_3. Here I again restrict the sample to only loans with an LTV ratio of 80. This restriction provides the same benefits as in column 2, but also strengthens the local randomness assumption on loan size. If borrowers are taking out a loan with the default LTV ratio of 80, their loan size is just 80% of price of the home, which is unlikely to be determined based on specified pool payups. Then, I restrict the sample to loans that are both near an LLPA credit score cutoff and near a loan size cutoff. I use county fixed effects and linearly control for loan size and credit score, allowing both relationships to vary across buckets. Here, I get point estimates of $\rho_1 = 0.85$, $\rho_2 = 0.36$, and $\rho_3 = 0.87$. We can reject the null hypothesis that $\rho_1 = \rho_2$ at the 1% level but not that $\rho_1 = \rho_3$. Pass-through of specified pool payups is different than that of TBA prices or LLPAs.

Column 5 estimates the same thing, but again loosens the sample restriction to any loan that is near a loan size cutoff and either within 5 of a credit score cutoff or within 2 of an LTV ratio cutoff. Point estimates are similar at $\rho_1 = 0.85$, $\rho_2 = 0.37$, and $\rho_3 = 0.89$. Again, we can only reject the null hypothesis that pass-through of TBA prices is equal to that of specified pool payups.

Finally, Column 6 estimates Equation 22. This makes use of all loans. I control for the continuous relationship between marginal costs and loan size, credit score, and LTV ratio by using (interacting) third degree polynomial terms. This specification assumes that these polynomial terms fully capture the relationship between these three running variables and the unobserved component of marginal cost. In exchange for this (strong) assumption, the regression can leverage additional data points away from the cutoffs, increasing power. I again include county fixed effects. Point estimates are qualitatively similar at $\rho_1 = 0.95$, $\rho_2 = 0.22$, and $\rho_3 = 0.82$. Here, only the null hypothesis that $\rho_1 = \rho_2$ can be rejected.

Online Appendix Section B demonstrates that the above results are robust to a wide array of specifications. Qualitatively similar results are obtained when one repeats the above analysis but uses borrower discount points or differences in TBA prices to calculate the price. It is worth noting that changing the notion of price used primarily changes the levels of the three pass-through rates rather than the magnitude of the difference between them. The results lead to the same conclusions when we use different regression discontinuity windows, estimate pass-through separately for each specified pool cut-off, or use a different base coupon rate for TBA prices. Online Appendix Section B contains the details of these robustness checks.

	(1) Present Value	(2) Present Value	(3) Present Value	(4) Present Value	(5) Present Value	(6) Present Value
TBA Price	0.85	0.97	0.99	0.85	0.85	0.95
	[0.74,0.96]	[0.85,1.10]	[0.87,1.11]	[0.75,0.94]	[0.74,0.96]	[0.80,1.10]
Payup	0.37			0.36	0.37	0.22
	[0.14,0.60]			[0.14,0.59]	[0.15,0.59]	[0.043,0.39]
LLPA		0.87	0.96	0.87	0.89	0.82
		[0.77,0.98]	[0.86,1.05]	[0.59,1.14]	[0.78,1.01]	[0.70,0.94]
Loan Size Controls		✓	✓			
Credit Controls	✓					
LTV Restriction	≤ 80	$= 80$	≤ 80	$= 80$	≤ 80	≤ 80
Sample Window	Loan Size	FICO	FICO or LTV	Size and FICO	Size and FICO or LTV	All
N	69,252	93,344	216,596	13,636	37,462	399,260
R^2	.652	.665	.67	.656	.653	.68
P(TBA = Spec)	.003			.001	.003	0
P(TBA = LLPA)		.195	.672	.858	.522	.175

Table 1: Borrower present values regressed on TBA prices, specified pool payups, and LLPAs. All regressions control for county fixed effects. All standard errors are clustered at the location, firm, time level. Point estimates for pass-through rates reported with 95% confidence intervals in brackets. Loan size controls are fixed effects for loan size calculated using \$50,000 buckets. Credit controls are fixed effects by LLPA bins. LTV restriction refers to whether the sample was restricted to borrowers with an LTV of 80 or not. Sample window refers to restriction of the sample to loans that are close to specified pool or LLPA cutoffs. Loan size refers to within \$5000 of a specified pool cutoff, FICO refers to within 5 of an LLPA credit score cutoff, and LTV refers to within 2 of an LLPA LTV cutoff. N reports the sample size. P(TBA = Spec) reports the p-value of an F-test comparing the coefficients (pass-through rates) on TBA prices and specified pool payups. P(TBA = LLPA) reports the p-value of an F-test comparing the coefficients on TBA prices and LLPAs.

5.5 Discussion

No matter the methodology used to convert interest rates into upfront prices, it appears that the pass-through rates of TBA prices and specified pool payups are different. It does not appear that the pass-through rates of TBA prices and LLPAs are different. This evidence is inconsistent with the models of equilibrium pricing found in the existing literature. In these models, all components of marginal cost should be passed through at the same rate. Here, each of the components of marginal cost - TBA price, specified pool payup, and LLPA - enter the firms' profit function in the same way. Each of these costs is common to all firms. Each is part of the same product ruling out differences in demand or firm conduct. Therefore, this evidence is inconsistent with these existing models.

On the other hand, this evidence is consistent with the model presented above if borrowers have less precise beliefs about specified pool payups than the other two components. As explained above, this seems like a reasonable explanation. Borrowers are likely aware of prevailing interest rates and their effect on the costs faced by lenders. Likewise, borrowers are likely to know that it is more costly to lend to borrowers with lower credit scores and lower down payments (higher LTV ratios). While they may not be aware of or understand the TBA market or the determination of LLPAs, they are likely to be aware of the correlations between marginal cost of lending and interest rates, credit scores, and LTV ratios. However, it seems unlikely that borrowers are aware of the specified pool market and their intuitions about the relationship between marginal cost and loan size would probably lead them in the opposite direction. The above evidence on pass-through rates rejects existing models but remains consistent with a model in which consumers' expectations impact pass-through.

5.6 Counterfactual Consumer Information

The above empirical exercises are primarily intended as a test of the novel theoretical prediction that pass-through can vary across components of marginal cost. However, the estimated pass-through rates are also sufficient statistics for measuring the gain in consumer welfare created by the specified pool market both under current conditions and under a counterfactual consumer information regime. These estimates allow us to use the model to predict the additional welfare that they could receive if they had the same information about specified pools as they do about general interest rates.

The specified pool market allows lenders to receive additional revenue from the origination of loans below a certain loan size. In this section, I study how much of the surplus generated by the specified pool market is received by consumers. Specifically, I use the model to perform the following thought experiment: holding all else fixed, how much surplus would a small loan size borrower lose if their loan were not eligible to be securitized as a specified pool (and had to be sold via TBA instead). To quantify the savings that these borrowers receive from the specified pool market, I estimate the upfront points they would need to pay the lender in order to maintain the same interest rate. This allows me to specify the total upfront savings a borrower receives from being eligible for their given specified pool. I can then compare these savings to the savings in a counterfactual in which consumer's information about the specified pool market is equally precise as their information about TBA prices.

With the above estimates of the pass-through rates, I can estimate these savings both under consumers' actual information and under the counterfactual. Without additional assumptions about consumers' expectations about specified pool payups, I cannot make any statements about the levels of price in the counterfactual information scenario.[19]

Table 2 reports the results of these counterfactual exercises for current borrowers that are eligible for each specified pool story. Throughout this section, I use the estimates from Table A.1, Column 1. These estimates allow me to express consumer surplus as the points the borrower would need to pay upfront to maintain the same interest rate if they were not eligible for the specified pool. One could easily use any of the above estimates to perform a similar exercise.

The first four rows present the results in points (dollars paid/saved upfront per hundred dollars of loan size). The first row reports the average payup received by lenders for borrowers whose loans were pooled in a given specified pool story. This is taken directly from the data and is generally decreasing in the loan size cutoff[20]. The second row reports estimated borrower savings due to specified pool eligibility (in points). Here we are comparing the price lenders would offer given the payup they actually received and the price they would offer given a payup of zero. Under the model, this can be calculated by simply multiplying the payup received by the pass-through rate on specified pool payups. Average savings for specified pool borrowers were between 0.25 and 0.61 points. The third row reports the estimated borrower savings due to specified pool eligibility in the counterfactual scenario in which borrowers' information about specified pool payups is equally precise as their information about TBA prices. In the model, by Corollary 2.1, two components of marginal cost have the same pass-through rate if and only if the precision of consumers' beliefs about the components are the same. We can therefore estimate the savings specified pool eligible borrowers would receive in this scenario by multiplying actual payup by the pass-through rate on TBA prices. In this scenario, we estimate that specified pool eligibility would save borrowers between 0.66 and 1.59 points. The fourth row reports the difference

[19]Depending on consumers' current expectations of payups, it is possible that prices for specified pool eligible borrowers decrease, prices for ineligible borrowers increase, or some combination of the two. Therefore, without additional assumptions, I can only make statements about the difference in surplus between otherwise equivalent eligible and ineligible borrowers. However, if one makes the reasonable assumption that borrowers currently believe payups are zero, then these results can be interpreted as the change in prices specified pool borrowers would receive in the counterfactual information setting.

[20]Average payups received by borrowers are higher for HHLB and 200k borrowers because of compositional effects. These stories were not widely traded at the beginning of the sample period, and became more common at a time when all payups were higher. On any given date, HHLB and 200k payups were less than LLB, MLB or HLB ones.

between these two estimates. This represents the additional savings specified pool eligible borrowers would receive if they had more precise information about payups. Estimates for these additional savings range from 0.41 to 0.98 points.

The middle four rows calculate savings per borrower. Here, I report the average loan size for each specified pool story and then each of the above savings estimates in dollars. I estimate that average savings created by the specified pool eligibility are between $415 and $561. With more precise information, these borrowers could have received an additional $671 to $905.

The final four rows report the total savings generated by the Fannie Mae specified pool market. I estimate that the current total borrower surplus generated is roughly $178 million dollars per year. I estimate that borrowers would save about $467 million dollars annually with counterfactual information. I therefore estimate that across these specified pool stories, borrowers miss out on $289 million dollars in savings per year.

	LLB	MLB	HLB	HHLB	200k
Specified Pool Payup in Points	1.48	1.32	0.77	0.84	0.62
Current Savings in Points	0.61	0.54	0.32	0.34	0.25
Counterfactual Savings in Points	1.59	1.42	0.83	0.90	0.66
Difference in Points	0.98	0.88	0.51	0.56	0.41
Loan Amount at Origination	68,016.16	98,710.89	131,258.26	162,615.29	188,333.34
Current Savings in Dollars	415.65	534.84	432.77	560.55	478.73
Counterfactual Savings in Dollars	1,086.59	1,398.19	1,131.35	1,465.39	1,251.49
Difference in Dollars	670.94	863.35	698.58	904.84	772.76
Fannie Mae Annual Issuance ($MM)	3,279.2	5,454.8	12,459.0	13,865.8	16,617.3
Annual Savings ($MM)	20.0	29.5	39.9	47.1	41.5
Annual Counterfactual Savings ($MM)	52.1	77.5	103.4	124.8	109.7
Annual Difference ($MM)	32.1	48.0	63.5	77.6	68.1

Table 2: Table reports averages for each variable grouped by specified pool story. Specified pool stories are LLB (Loan Size $\leq$ $85,000), MLB ($85,000 $<$ Loan Size $\leq$ $110,000), HLB ($110,000 $<$ Loan Size $\leq$ $150,000), HHLB ($150,000 $<$ Loan Size $\leq$ $175,000), and 200k ($175,000 $<$ Loan Size $\leq$ $200,000). All measures of savings are calculated using estimates from Table A.1, column 1. Each is reported in points (dollars up-front per $100 in loan amount), upfront dollars (points/100 multiplied by loan size), and total annual savings (points/100 multiplied by average annual Fannie Mae pool issuance). Average annual pool issuance for Fannie Mae by story is calculated from Fannie Mae's MBS Data Dynamics pool details data.

6 Conclusion

This paper provides a model of pass-through with consumer search and cost uncertainty which is able to explain a wide variety of empirical patterns in pass-through, holding fixed standard determinants of pass-through. I find that consumers' expectations about the costs facing firms are critical in determining whether firms can pass these costs on to consumers. Generally, firms are more able to pass through costs that consumers are aware of. Pass-through is higher in environments where consumers have precise information about the marginal cost of production. If firms are able to reveal their production costs (at a cost), this can lead to asymmetric pass-through in which firms choose to reveal higher costs, leading to high pass-through of higher costs. I find that pass-through rates differ across components of marginal cost in the US mortgage market, a fact that is consistent with this model, and inconsistent with existing theory.

This research has many additional implications. For macroeconomists, this model may serve as a new mechanism for explaining the amplification or dampening of inflation. If consumers in an economy believe that inflation is high and that firms are facing unusually high production costs, this could lead to firms raising prices. It would be interesting to see this model applied to a model of the aggregate macroeconomy.

For microeconomists, this research has important implications for counterfactuals made based on measured pass-through rates. If consumers have different beliefs over different components of the marginal cost of production, these components could have different associated pass-through rates. Therefore, a pass-through rate measured using one element of marginal cost may not be applicable to another.

The model also has implications for tax incidence that might be of interest to policy makers and public economists. The model implies that taxes will have different effects on consumer welfare depending on the information consumers have about the tax. This could break the equivalence of statutory incidence. A non-salient tax that is placed on producers may have lower pass-through to prices than a similar tax placed on consumers. [21]

For regulators, this model adds an additional wrinkle to discussions over whether more accurate information will help or harm consumers. Here, the effect is ambiguous and depends on the directionality of underlying costs, as well as the distribution of these costs

[21] Note that this operates through a different channel and in the opposite direction than the results of Chetty, Looney, and Kroft (2009). In their model, consumers may not fully consider all of the components of the offered price. In this model, consumers perfectly understand prices, but must use their beliefs about costs to form expectations about other prices in the market.

across a population. Market regulators should note that pass-through is dependent on both the level of competition in a market and consumers' beliefs about these costs.

References

AGARWAL, S., I. BEN-DAVID, AND V. YAO (2017): "Systematic mistakes in the mortgage market and lack of financial sophistication," *Journal of Financial Economics*, 123(1), 42–58.

AGARWAL, S., S. CHOMSISENGPHET, N. MAHONEY, AND J. STROEBEL (2014): "A simple framework for estimating consumer benefits from regulating hidden fees," *The Journal of Legal Studies*, 43(S2), S239–S252.

——— (2015): "Regulating consumer financial products: Evidence from credit cards," *The Quarterly Journal of Economics*, 130(1), 111–164.

ALEXANDROV, A., AND S. KOULAYEV (2018): "No shopping in the us mortgage market: Direct and strategic effects of providing information," *Consumer Financial Protection Bureau Office of Research Working Paper*, (2017-01).

ALLEN, J., R. CLARK, AND J.-F. HOUDE (2014): "The effect of mergers in search markets: Evidence from the Canadian mortgage industry," *American Economic Review*, 104(10), 3365–96.

ANAGOL, S., AND H. H. KIM (2012): "The impact of shrouded fees: Evidence from a natural experiment in the Indian mutual funds market," *American Economic Review*, 102(1), 576–593.

BENABOU, R., AND R. GERTNER (1993): "Search with learning from prices: does increased inflationary uncertainty lead to higher markups?," *The Review of Economic Studies*, 60(1), 69–93.

BENETTON, M., A. GAVAZZA, AND P. SURICO (2021): "Mortgage pricing and monetary policy," .

BHUTTA, N., A. FUSTER, AND A. HIZMO (2020): "Paying too much? Price dispersion in the US mortgage market," .

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2013): "Salience and consumer choice," *Journal of Political Economy*, 121(5), 803–843.

——— (2015): "Competition for Attention," *The Review of Economic Studies*, 83(2), 481–513.

BORENSTEIN, S., A. C. CAMERON, AND R. GILBERT (1997): "Do gasoline prices respond asymmetrically to crude oil price changes?," *The Quarterly Journal of Economics*, 112(1), 305–339.

BULOW, J. I., AND P. PFLEIDERER (1983): "A note on the effect of cost changes on prices," *Journal of Political Economy*, 91(1), 182–185.

BURDETT, K., AND K. L. JUDD (1983): "Equilibrium price dispersion," *Econometrica: Journal of the Econometric Society*, pp. 955–969.

BUSSE, M., J. SILVA-RISSO, AND F. ZETTELMEYER (2006): "$1,000 cash back: The pass-through of auto manufacturer promotions," *American Economic Review*, 96(4), 1253–1270.

CABRAL, L., AND S. GILBUKH (2020): "Rational buyers search when prices increase," *Journal of Economic Theory*, 187, 104998.

CATTANEO, M. D., AND R. TITIUNIK (2022): "Regression discontinuity designs," *Annual Review of Economics*, 14, 821–851.

CHETTY, R. (2009): "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 1(1), 451–488.

CHETTY, R., A. LOONEY, AND K. KROFT (2009): "Salience and taxation: Theory and evidence," *American Economic Review*, 99(4), 1145–77.

DANA JR, J. D. (1994): "Learning in an equilibrium search model," *International Economic Review*, pp. 745–771.

DHOLAKIA, U. M. (2021): "If you're going to raise prices, tell customers why," *Harvard Business Review*.

DIAMOND, P. A. (1971): "A model of price adjustment," *Journal of economic theory*, 3(2), 156–168.

D'ACUNTO, F., U. MALMENDIER, J. OSPINA, AND M. WEBER (2021): "Exposure to grocery prices and inflation expectations," *Journal of Political Economy*, 129(5), 1615–1639.

EMMER, M. (2022): "How To Pass On Price Increases Without Alienating Customers," *Forbes*.

EYSTER, E., K. MADARÁSZ, AND P. MICHAILLAT (2021): "Pricing under fairness concerns," *Journal of the European Economic Association*, 19(3), 1853–1898.

FISHMAN, A. (1996): "Search with learning and price adjustment dynamics," *The Quarterly Journal of Economics*, 111(1), 253–268.

FUSTER, A., L. S. GOODMAN, D. O. LUCCA, L. MADAR, L. MOLLOY, AND P. WILLEN (2013): "The rising gap between primary and secondary mortgage rates," *Economic*

Policy Review, 19(2).

FUSTER, A., S. H. LO, AND P. S. WILLEN (2017): “The time-varying price of financial intermediation in the mortgage market,” Discussion paper, National Bureau of Economic Research.

FUSTER, A., D. LUCCA, AND J. VICKERY (2023): “15. Mortgage-backed securities1,” *Research Handbook of Financial Markets*, p. 331.

GOLDBERG, P. K., AND M. M. KNETTER (1997): “Goods Prices and Exchange Rates: What Have We Learned?,” *Journal of Economic Literature*, 35(3), 1243–1272.

HEASLIP, E. (2022): “Inflation Driving Up Prices? How to Communicate a Price Increase to Customers,” *CO— by U.S. Chamber of Commerce*.

JAFFE, S., AND E. G. WEYL (2013): “The first-order approach to merger analysis,” *American Economic Journal: Microeconomics*, 5(4), 188–218.

JANSSEN, M., P. PICHLER, AND S. WEIDENHOLZER (2011): “Oligopolistic markets with sequential search and production cost uncertainty,” *The RAND Journal of Economics*, 42(3), 444–470.

JANSSEN, M., AND S. SHELEGIA (2020): “Beliefs and consumer search in a vertical industry,” *Journal of the European Economic Association*, 18(5), 2359–2393.

KAPLAN, G., AND G. MENZIO (2015): “The morphology of price dispersion,” *International Economic Review*, 56(4), 1165–1206.

KROFT, K., J.-W. LALIBERTÉ, R. LEAL-VIZCAÍNO, AND M. J. NOTOWIDIGDO (2023): “Salience and Taxation with Imperfect Competition,” *The Review of Economic Studies*, p. rdad028.

LIU, L. (2019): “Non-salient fees in the mortgage market,” *Available at SSRN 3437374*.

MARION, J., AND E. MUEHLEGGER (2011): “Fuel tax incidence and supply conditions,” *Journal of public economics*, 95(9-10), 1202–1212.

MILLER, N. H., M. REMER, C. RYAN, AND G. SHEU (2016): “Pass-through and the prediction of merger price effects,” *The Journal of Industrial Economics*, 64(4), 683–709.

MILLER, N. H., M. REMER, AND G. SHEU (2013): “Using cost pass-through to calibrate demand,” *Economics Letters*, 118(3), 451–454.

MONTAG, F., A. SAGIMULDINA, AND M. SCHNITZER (2021): “Does tax policy work when consumers have imperfect price information? Theory and evidence,” .

MORAGA-GONZÁLEZ, J. L., Z. SÁNDOR, AND M. R. WILDENBEEST (2023): “Consumer search and prices in the automobile market,” *The Review of Economic Studies*, 90(3), 1394–1440.

MUEHLEGGER, E., AND R. L. SWEENEY (2021): "Pass-through of own and rival cost shocks: Evidence from the US Fracking boom," *The Review of Economics and Statistics*, pp. 1–32.

NAKAMURA, E., AND J. STEINSSON (2011): "Price setting in forward-looking customer markets," *Journal of Monetary Economics*, 58(3), 220–233.

NAKAMURA, E., AND D. ZEROM (2010): "Accounting for incomplete pass-through," *The review of economic studies*, 77(3), 1192–1230.

PELTZMAN, S. (2000): "Prices rise faster than they fall," *Journal of political economy*, 108(3), 466–502.

PLESS, J., AND A. A. VAN BENTHEM (2019): "Pass-through as a test for market power: An application to solar subsidies," *American Economic Journal: Applied Economics*, 11(4), 367–401.

ROTEMBERG, J. J. (2005): "Customer anger at price increases, changes in the frequency of price adjustment and monetary policy," *Journal of monetary economics*, 52(4), 829–852.

SCHARFSTEIN, D., AND A. SUNDERAM (2016): "Market power in mortgage lending and the transmission of monetary policy," *Unpublished working paper. Harvard University*.

STAHL, D. O. (1989): "Oligopolistic pricing with sequential consumer search," *The American Economic Review*, pp. 700–712.

SUMNER, D. A. (1981): "Measurement of monopoly behavior: an application to the cigarette industry," *Journal of Political Economy*, 89(5), 1010–1019.

TAPPATA, M. (2009): "Rockets and feathers: Understanding asymmetric pricing," *The RAND Journal of Economics*, 40(4), 673–687.

WEBER, M., F. D'ACUNTO, Y. GORODNICHENKO, AND O. COIBION (2022): "The Subjective Inflation Expectations of Households and Firms: Measurement, Determinants, and Implications," *Journal of Economic Perspectives*, 36(3), 157–84.

WEYL, E. G., AND M. FABINGER (2013): "Pass-through as an economic tool: Principles of incidence under imperfect competition," *Journal of Political Economy*, 121(3), 528–583.

YANG, H., AND L. YE (2008): "Search with learning: understanding asymmetric price adjustments," *The RAND Journal of Economics*, 39(2), 547–564.

Online Appendix A Proofs

A.1 Proposition 1

Equilibrium existence (Lemma 1) is proven in Stahl (1989).

To prove that pass-through rate is equal to 1, consider a normalization of the problem in which $\hat{p} = p - C$ and $\hat{C} = c - c = 0$. Given these new variables, there exists a function $\hat{F}$ such that, $\hat{F}(\hat{p}) = \hat{F}(p - C) = F(p)$, where F is defined as in Equation 4. Then the firms profit function (Equation 2) can be rewritten in terms of $\hat{F}$:

$$\hat{\pi}(\hat{p}) = (\frac{1-\mu}{N} + \mu(1 - \hat{F}(p))^{N-1})\hat{p}$$

The set of markups, $\hat{p}$, that maximize $\hat{\pi}(\hat{p})$, do not depend on C.

We can also rewrite the consumers' search indifference condition in terms of these normalized variables:

$$\hat{r} + C = s + \int_{\underline{\hat{p}}+C}^{\hat{r}+C} [f(\hat{p} + C)\hat{p} + C] dp_i$$

Where $\hat{r} = r - C$ and $\underline{\hat{p}} = \underline{p} - C$. One can subtract C from both sides, substitute $\hat{f}$ for f, and use a change of variable from p to $\hat{p}$ (and adjust the bounds of integration accordingly):

$$\hat{r} = s + \int_{\underline{\hat{p}}}^{\hat{r}} [\hat{f}(\hat{p})\hat{p}] d\hat{p}$$

The markup at which consumers are indifferent between searching and not searching, $\hat{r}$, therefore does not depend on C. Equilibrium is then defined in terms of the firms' strategic object $\hat{p}$ and the consumers' strategic object $\hat{r}$, neither of which depend on C. As none of the equilibrium strategies in markups depend on C, the average chosen (or transacted) markup will not depend on C. The average transacted price is then, simply, the average transacted markup plus C. Average transacted markup does not depend on C so the derivative of average transacted price and C must be 1.

That all components of cost are passed through at a rate of 1 follows simply from the previous statement and that $C = \sum_{j=1}^{J} c_j$.

A.2 Proposition 2

To prove existence of this equilibrium it will first be useful to characterize consumers' returns from search. Assume that there exists a reservation price strategy such that consumers purchase a good if they receive a price less than or equal to r but search otherwise. Then, from the previous section, firms' prices will be found on the interval $[\underline{p}, r]$ and distributed according to the density function $f(p|C)$ where:

$$\underline{p} = \frac{(1-\mu)r + 2\mu c}{1-\mu+2\mu}.$$

and

$$f(p|C) = \frac{(1-\mu)(r-C)}{2\mu(p-C)^2},$$

Consumers with a single price quote (non-shoppers) must decide whether to search again and do so if the expected value of searching is positive. This value of searching depends on the first price quote, p, the consumer received and their prior beliefs. Given p, one can use Bayes' rule to compute the posterior probability of the firms' marginal costs:

$$h(C|X,p) = \frac{h(C|X)f(p|C)}{\int h(\gamma|X)f(p|\gamma)d\gamma}$$

Plugging in the equilibrium density function for prices:

$$h(C|x,p) = \frac{h(C|X)\frac{(1-\mu)(r-C)}{2\mu(p-C)^2}}{\int h(\gamma|X)\frac{(1-\mu)(r-\gamma)}{2\mu(p-\gamma)^2}d\gamma}$$

Critically, $h(C|x,p) = 0$ for all $C > p$, as this would result in negative profits for the firm and will never be played in equilibrium. Given these updated beliefs, the consumers returns to search are:

$$\phi(p) = -s + p - \int_c h(C|X,p) * \int_{\underline{p}}^{r} f(q|C) * q dq dC$$

The second integral is the average offered price given C and can be evaluated for any C:

$$\int_{\underline{p}}^{r} \frac{(1-\mu)(r-C)}{2\mu(q-C)^2} * q = \frac{(1-\mu)(r-C)}{2\mu}[\log(q-C) - \frac{C}{q-C}]|^{r}_{\frac{(1-\mu)r+2\mu C}{1-\mu+2\mu}}$$

$$= \frac{(1-\mu)(r-C)}{2\mu}[\log(\frac{r-C}{\frac{(1-\mu)r+2\mu C}{1-\mu+2\mu}-C}) - \frac{C}{r-C} + \frac{C}{\frac{(1-\mu)r+2\mu C}{1-\mu+2\mu}-C}]$$

Rearranging:

$$= \frac{(1-\mu)(r-C)}{2\mu}[\log(\frac{1+\mu}{1-\mu})) + (\frac{2\mu}{1-\mu})(\frac{C}{r-C})]$$

$$\int_{\underline{p}}^{r} f(q|C) * q dq = (1-A(\mu))r + A(\mu)C$$

$A(\mu)$ is a constant that takes on values between 0 and 1 and is strictly increasing in μ, the number of shoppers. So the average offered price is a convex combination of r and C. We can then plug this back into the consumer search value function:

$$\phi(p) = -s + p - \int_{c} h(C|X,p) * [(1-A(\mu))r + A(\mu)C]dC$$

Then, noting that r does not depend on the actual draw of C, it only depends on the signals the consumer has:

$$\phi(p) = -s + p - (1-A(\mu))r - A(\mu)\int_{C} h(C|X,p) * CdC$$

For this function to be monotonically increasing in p, it must be that, for all p:

$$p > A(\mu)\int_{C} h(C|X,p) * CdC = A(\mu)E_h[C|X,p]$$

This statement will always be true. Given $h(C|X,p) = 0$ for all $C > p$, it must be that $E_h[C|X,p] < p$. $A(\mu) \leq 1$, therefore this statement will always hold. Therefore, conditional on firms playing the equilibrium price distribution for a reservation price strategy of an arbitrary r, consumers' value of search is increasing in the price that they see. Therefore, this function must have a root r^*, such that above r^*, the value of search is positive, and below r^* the value of search is negative. The search value function can be rewritten in terms of the signal X and the expected noise ϵ and then we can solve for the root in which $p = r = r^*$ which will give us Equation 9:

$$r^* = X + \frac{s}{A(\mu)} + E[\epsilon|p = r^*]$$

Given the construction of this value r^*, it is optimal for consumers to search when the first price they see is $p > r$ and purchase when $p \leq r$.

Next, one can show that the distribution of prices played by firms is optimal. It is clear from the construction of the distribution that profits from playing any price in $[\underline{p}, r^*]$ are the same, thus firms are indifferent between playing any of these prices. It is also obvious that they would not want to play any price less than $\underline{p}$ because $1 - F(\underline{p}) = 1$, and they are already maximizing their market share. Therefore playing a lower price can only lead to lower profits. Firms will not play prices above r^* because this will induce the consumer to search. If the consumer searches, the other firm will be playing a price less than r^* in equilibrium, therefore the firm will earn zero profits. All that remains to be shown is that the profits from playing a price in $[\underline{p}, r^*]$, which are equal to the profits from playing r^*, are greater than or equal to zero. For this to be true, it must be that $r^* \geq c$ for all possible C. Therefore for this equilibrium to exist, it must be that for all possible ϵ:

$$r^* = X + \frac{s}{A(\mu)} + E[\epsilon|p = r^*] \geq X + \epsilon = C$$

Simplifying this expression, there must be an upper bound $\bar{\epsilon}$ such that $h(\epsilon) > 0 \implies \epsilon \leq \bar{\epsilon}$ and

$$\bar{\epsilon} \leq \frac{s}{A(\mu)} + E[\epsilon|p = r^*]$$

If this restriction is met, then the firms' strategy is also optimal, and the equilibrium exists.

Given this equilibrium, it is straightforward to show that the pass-through rate of costs captured by consumers beliefs (X) is 1 and the pass-through rate on costs that are not (ϵ) is μ. One can plug this reservation price into the expression for the density of prices, then use this to get that the average transacted price is:

$$\bar{p} = X + \mu\epsilon + (1 - \mu)(\hat{\epsilon} + \frac{s}{A(\mu)})$$

Clearly the derivative of $\bar{p}$ with respect to X is 1 and the derivative of $\bar{p}$ with respect to ϵ is μ.

To demonstrate the second part of this proposition note that $c_j = x_j + \epsilon_j$. From the above equation, pass-through of a component of cost c_j is a convex combination of 1 and μ where the weight placed on 1 is $\frac{\partial x_j}{\partial c_j}$ and the weight placed on μ is 1 minus this value or $\frac{\partial \epsilon_j}{\partial c_j}$. As μ

is constant across cost components, if $\frac{\partial x_1}{\partial c_1} = \frac{\partial x_2}{\partial c_2} \iff \rho_1 = \rho_2$.

A.3 Proposition 3

To demonstrate this, I first need to show how individuals beliefs evolve. At time t, the individual expects the cost index to be $\hat{c}^{\dagger}_{t|t-1} = \hat{c}_{t-1}$ and the variance of this is $\hat{P}_{t|t-1} = \hat{P}_{t-1} + \sigma^2_\nu$. The individual updates their beliefs using the Kalman gain term:

$$K_t = \frac{\hat{P}_{t-1} + \sigma^2_\nu}{\hat{P}_{t-1} + \sigma^2_\nu + \sigma^2_\xi} < 1. \tag{23}$$

This term captures the degree to which consumers update their beliefs upon receiving new information. We can use this to solve for the steady state variance of the consumers' estimate of the untransformed production cost:

$$\hat{P} = \frac{\sqrt{4\sigma^2_\xi \sigma^2_\nu + (\sigma^2_\nu)^2} - \sigma^2_\nu}{2} \tag{24}$$

The variance of consumers' information is increasing in the variance of the signal (σ^2_ξ) and the variance of the shocks to the underlying untransformed cost (σ^2_ν). It is worth noting that if *either* consumers perfectly observe costs ($\sigma^2_\xi = 0$) or costs evolve according to a deterministic process ($\sigma^2_\nu = 0$), than this variance is zero, and the model simplifies to the perfect information Stahl (1989) case. We can also plug this steady state variance into Equation 23 to get the steady state Kalman gain term $\hat{K}$ which is less than one and is decreasing in σ^2_ξ.

This defines the signal structure and updating process of the untransformed cost. Assume that the consumers have observed signals for a sufficient period of time and thus believe that costs are normally distributed with mean $\hat{c}^{\dagger}_{t-1} = 0$ and variance $\hat{P}$ ($\hat{c}^{\dagger}_{t-1}$ can be at any level, but I will assume it is equal to zero to simplify the algebra). We will assume that consumers' mean beliefs are correct at the time of analysis and thus $c^{\dagger}_{t-1} = 0$ and $c_{t-1} = \bar{c}$. Denote the inverse of t as $s : (-\gamma, \gamma) \to \mathbb{R}$.

Given these prior beliefs about the cost index, the consumers believe that actual marginal costs are distributed over the interval $(\bar{c}-\gamma, \bar{c}+\gamma)$ according to distribution function:

$$G(c_{t-1}) = F(\frac{s(c_{t-1}-\bar{c})}{\hat{P}}) \tag{25}$$

Where F is the standard normal distribution. Beliefs over costs will always be bounded and satisfy Assumption 1 so the average price for any realization of cost (c) and noise (ξ) is defined by Equation 10 which can be rearranged slightly as follows:

$$\bar{p}(c,\xi) = x(c,\xi) + \mu(c - x(c,\xi)) + (1-\mu)(\hat{\epsilon}(c,\xi) + \frac{s_i}{A(\mu)}) \tag{26}$$

Here I have rewritten ϵ as $c - x$ (the actual realization of the cost less the consumers' expectation of that cost). I have also expressed the mean of consumers' beliefs over costs (x) and the mean of consumers beliefs over ϵ, conditional on seeing a price of r ($\hat{\epsilon}$) as functions of c and ξ. Both of these values are actually just functions of the signal that the consumer received y. However y can be decomposed into the actual cost and the noise term ξ. This decomposition will be useful when we integrate over possible signals. The average price given cost c is then the average price across all possible realizations of ξ:

$$\bar{p}(c) = \int_{-\infty}^{\infty} [x(c,\xi) + \mu(c - x(c,\xi)) + (1-\mu)(\hat{\epsilon}(c,\xi) + \frac{s_i}{A(\mu)})]g(\xi)d\xi \tag{27}$$

Where $g(\xi)$ is normal with mean 0, variance σ_ξ^2. Average pass-through is then the derivative of this average price with respect to cost:

$$\rho(c) = \int_{-\infty}^{\infty} [x_1'(c,\xi) + \mu(1 - x_1'(c,\xi)) + (1-\mu)\hat{\epsilon}_1'(c,\xi)]g(\xi)d\xi = \mu + (1-\mu)\int_{-\infty}^{\infty} [x_1'(c,\xi) + \hat{\epsilon}_1'(c,\xi)]g(\xi)d\xi \tag{28}$$

To understand this term we need the expressions for $x(c,\xi)$ and $\hat{\epsilon}(c,\xi)$. I will show that at $c = \bar{c}$, $0 \leq x_1'(\bar{c},\xi) + \hat{\epsilon}_1'(\bar{c},\xi) \leq 1$ and thus $\mu \leq \rho(\bar{c}) \leq \mu + \hat{K}(1-\mu) \leq 1$.

First, we know that given our assumed prior $\hat{c}_{t-1} = \bar{c}$, a new cost c, and a noise shock ξ, consumers' will believe that the cost index $c^\dagger$ is distributed normally with mean

$$\hat{c}^\dagger(c,\xi) = \hat{K} * (s(c-\bar{c}) + \xi) \tag{29}$$

and variance $\hat{P}$. They therefore believe that the mean (transformed) cost is:

$$x(c, \xi) = \bar{c} + \int_{-\infty}^{\infty} t(\hat{K} * (s(c - \bar{c}) + \xi) + \zeta) f(\zeta) d\zeta \tag{30}$$

Where $f(\zeta)$ is normal mean 0, variance $\hat{P}$. The derivative of this expression with respect to c is then:

$$x_1'(c, \xi) = \hat{K} \int_{-\infty}^{\infty} t'(\hat{K} * (s(c - \bar{c}) + \xi) + \zeta) * s'(c - \bar{c}) g(\zeta) d\zeta = \hat{K} \frac{\int_{-\infty}^{\infty} t'(\hat{K} * (s(c - \bar{c}) + \xi) + \zeta)}{t'(s(c - \bar{c}))} g(\zeta) d\zeta \tag{31}$$

Under no transformation, t is linear, t' is constant, and this derivative is simply $\hat{K}$. This expression is generally difficult to evaluate, however, at $c = \bar{c}$, $t'(s(0))$ is at its maximum (we assumed 0 is the inflection point of t). Therefore at $c = \bar{c}$, the denominator is weakly greater than the numerator for all values of ζ being integrated over. Furthermore, $t' > 0$ at all points so we know that this fraction term lies between zero and one and thus $0 \leq x_1'(\bar{c}, \xi) \leq \hat{K} \leq 1$.

When there is no transformation, Assumption 2 is met and $\hat{\epsilon}_1'(c, \xi) = 0$. This would result in a pass-through rate of $\mu + \hat{K}(1 - \mu)$ which lies between μ and 1. However, given the curvature of t, we need to account for the $\hat{\epsilon}_1'(c, \xi)$ term. $\hat{\epsilon}$ is defined implicitly and will generally be difficult to define analytically. However, we can prove that $-x_1'(\bar{c}, \xi) \leq \hat{\epsilon}_1'(\bar{c}, \xi) \leq 0$, which would complete the proof that $\mu \leq \rho(\bar{c}) \leq 1$.

To show this, we can define $\hat{\epsilon}(y)$ where $y = s(c - \bar{c} + \xi)$. Then:

$$\hat{\epsilon}(y) = E[c - E[c|y = y]|y = y, p = r] = E[c|y = y, p = r] - E[c|y = y] \tag{32}$$

The probability of seeing a price of r (conditional on y) is strictly increasing in c. Therefore $E[c|y = y, p = r] \geq E[c|y = y]$ and $\hat{\epsilon}(y) > 0$. Taking the derivative of $\hat{\epsilon}$ with respect to y:

$$\hat{\epsilon}'(y) = \frac{\partial E[c|y = y, p = r]}{\partial y} - \frac{\partial E[c|y = y]}{\partial y} = \frac{\partial E[c|y = y, p = r]}{\partial y} - x_1'(\bar{c}, \xi) \tag{33}$$

The first term must be smaller than the second term. Fixing $p = r$ narrows the prior distribution over costs relative to not knowing p. Therefore, the signal y causes a smaller update to the expected cost when we fix the price to r. If $0 < \frac{\partial E[c|y=y,p=r]}{\partial y} < x_1'(\bar{c}, \xi)$, then

$-x_1'(\bar{c},\xi) \leq \hat{\epsilon}_1'(\bar{c},\xi) \leq 0$. This implies $\mu \leq \rho(\bar{c}) \leq 1$.

To see that $\rho(\bar{c})$ is decreasing in σ_ξ, it is useful to rewrite the expression for $\rho(\bar{c})$ explicitly in terms of σ_ξ:

$$\rho(c) = \mu + (1-\mu)\int_{-\infty}^{\infty}[x_1'(c,\xi*\sigma_\xi^2,\sigma_\xi^2) + \hat{\epsilon}_1'(c,\xi*\sigma_\xi^2,\sigma_\xi^2)]\phi(\xi)d\xi \tag{34}$$

Where $\phi(\xi)$ is the standard normal density and the third argument of each function is the variance of the noise term (which enters $\hat{K}$). Then taking the derivative of this function with respect to σ_ξ^2 we get:

$$\frac{\partial\rho(\bar{c})}{\partial\sigma_\xi^2} = (1-\mu)\int_{-\infty}^{\infty}[\xi*(x_{12}''(c,\xi*\sigma_\xi^2,\sigma_\xi^2)+\hat{\epsilon}_{12}''(c,\xi*\sigma_\xi^2,\sigma_\xi^2))+x_{13}''(c,\xi*\sigma_\xi^2,\sigma_\xi^2)+\hat{\epsilon}_{13}''(c,\xi*\sigma_\xi^2,\sigma_\xi^2)]\phi(\xi)d\xi \tag{35}$$

Under no transformation, the first two terms (the ones multiplied by ξ) are equal to zero. The fourth term would also be zero because $\hat{\epsilon}$ does not depend on c. The third term would simply be $\frac{\partial\hat{K}}{\partial\sigma_\xi^2}$ which is negative. Therefore, if t is linear, $\frac{\partial\rho(\bar{c})}{\partial\sigma_\xi^2} < 0$. However, with our transformation we need to account for distortions caused by t.

From the identity in Equation 32, we know that $x(c,\xi*\sigma_\xi^2,\sigma_\xi^2) + \hat{\epsilon}(c,\xi*\sigma_\xi^2,\sigma_\xi^2) = E[c|y = y, p = r]$. Denote $\hat{c}_r = E[c|y = y, p = r]$ to capture the consumers' expectation of firms costs conditional on the signal they see and seeing a price of r. Then:

$$\frac{\partial\rho(\bar{c})}{\partial\sigma_\xi^2} = (1-\mu)\int_{-\infty}^{\infty}[\xi*\hat{c}_{r12}''(c,\xi*\sigma_\xi^2,\sigma_\xi^2) + \hat{c}_{r13}'''(c,\xi*\sigma_\xi^2,\sigma_\xi^2)]\phi(\xi)d\xi \tag{36}$$

This expectation $\hat{c}_r$ is a function of the signal y which is a function of c and ξ. When ξ is positive (and $c = 0$), $\hat{c}_{r12}''$ will be negative. As ξ increases and moves away from zero, t' decreases which causes the consumer to place less probability weight on these higher realizations of c. Likewise when ξ is negative, $\hat{c}_{r12}''$ will be positive because, as ξ decreases from zero, t becomes flatter, and updating is dampened. Therefore the term $\xi\hat{c}_{r12}''$ will be weakly negative for all ξ.

The last term is clearly negative. σ_ξ^2 enters $\hat{c}_{r1}'$ only through $\hat{K}$. $\hat{K}$ is decreasing in the variance of the signal so this term is negative. Therefore both terms of the integral are negative for all ξ and the whole integral will be negative. Then $\rho(\bar{c})$ is decreasing in σ_ξ^2.

Finally, $\rho(\bar{c})$ is clearly decreasing in μ. Equation 34 shows that pass-through is a convex combination of 1 and a term less than one (the integral) where μ determines how much weight is placed on 1. Increasing μ increases $\rho(\bar{c})$ and moves pass-through closer to one.

A.4 Corollary 1

That the probability of pass-through greater than one is zero when $\sigma_\xi^2 = 0$ is a direct result of Proposition 1. The remainder of the proof refers to the case of $\sigma_\xi^2 > 0$.

Here we are again assuming that $\hat{c}_{t-1} = c_{t-1} = \bar{c}$. In the previous proof, I show that average transacted price for a given signal and underlying cost can be written as:

$$\bar{p}(c, \xi, \sigma_\xi^2) = x(c, \xi, \sigma_\xi^2) + \mu(c - x(c, \xi, \sigma_\xi^2)) + (1 - \mu)(\hat{\epsilon}(c, \xi, \sigma_\xi^2) + \frac{s_i}{A(\mu)}) \tag{37}$$

Assume that cost has increased by $\delta > 0$. Then the change in price can be written as a function of ξ (and σ_ξ):

$$\Delta\bar{p}(\xi, \sigma_\xi^2) = \mu\delta + (1 - \mu)[\Delta x(\bar{c} + \delta, \xi, \sigma_\xi^2) + \Delta\hat{\epsilon}(\bar{c} + \delta, \xi, \sigma_\xi^2)] \tag{38}$$

We are interested in finding the probability that $\Delta\bar{p}(\xi, \sigma_\xi^2) > \delta$. I will show that there is a cutoff value ξ^* above which $\Delta\bar{p}(\xi, \sigma_\xi^2) > \delta$ and below which $\Delta\bar{p}(\xi, \sigma_\xi^2) < \delta$. First, writing out the expression for when observed pass-through will be greater than one:

$$\delta \leq \mu\delta + (1 - \mu)[\Delta x(\bar{c} + \delta, \xi, \sigma_\xi^2) + \Delta\hat{\epsilon}(\bar{c} + \delta, \xi, \sigma_\xi^2)] \tag{39}$$

Rearranging:

$$\delta \leq \Delta x(\bar{c} + \delta, \xi, \sigma_\xi^2) + \Delta\hat{\epsilon}(\bar{c} + \delta, \xi, \sigma_\xi^2) = \Delta\hat{c}_r(\bar{c} + \delta, \xi, \sigma_\xi^2) \tag{40}$$

The left hand side of this inequality is a constant. The right hand side is monotonically increasing in ξ (it is $\Delta\hat{c}_r$ where $\hat{c}_r$ is once again the expected cost upon seeing a price of r). Consumers' expectation of cost conditional on seeing a price of r is strictly increasing in the signal that they see, which is increasing in ξ. Therefore, if there exists a value ξ^* of ξ for which this inequality binds, it must be that the inequality is satisfied if and only if

$\xi \geq \xi^*$. We can write this as an equality in terms of ξ^*:

$$\delta = \Delta \hat{c}_r(\bar{c} + \delta, \xi, \sigma_\xi^2) \tag{41}$$

We can decompose the right hand side of the equation into the change in consumers' beliefs in the un-transformed cost case and a residual term that captures the distortion introduced by the transformation:

$$\delta = \hat{K}(\delta + \xi^*) + \zeta(\delta, \xi^*, \sigma_\xi^2) \tag{42}$$

The function ζ takes on a negative value and is decreasing (becoming more negative) in its argument ξ. As $\hat{c}_r$ is strictly positive for positive $\delta + \xi$, ζ is also bound between $\hat{K}(\delta + \xi^*)$ and 0. We can then rearrange the above equation to implicitly solve for ξ^*:

$$\xi^* = \frac{\sigma_\xi^2 \delta}{\hat{P} + \sigma_\nu^2} - \frac{\zeta(\delta, \xi^*, \sigma_\xi^2)}{\hat{K}} \tag{43}$$

As ζ is negative, ξ^* takes on a greater value than it does in the un-transformed cost case. When there is no transformation, this results in the expression in the main text.

The probability that a pass-through rate greater than one is observed is:

$$Pr(\hat{\rho} > 1) = Pr(\xi > \xi^*) = 1 - \Phi(\frac{\xi^*}{\sigma_\xi}) \tag{44}$$

Where Φ is the standard normal distribution. Therefore, the probability of a pass-through rate greater than one is positive if $\frac{\xi^*}{\sigma_\xi}$ is finite.

$$\frac{\xi^*}{\sigma_\xi} = \frac{\sigma_\xi \delta}{\hat{P} + \sigma_\nu^2} - \frac{\zeta(\delta, \xi^*, \sigma_\xi^2)}{\hat{K}\sigma_\xi} \tag{45}$$

The first term, is clearly positive and finite for all $\sigma_\xi \in (0, \infty)$. The term $\zeta(\delta, \xi^*, \sigma_\xi^2)$ is bounded, therefore $-\frac{\zeta(\delta, \xi^*, \sigma_\xi^2)}{\hat{K}}$ is positive and finite for all $\sigma_\xi \in (0, \infty)$. Therefore, $Pr(\hat{\rho} > 1) > 0$.

When $\sigma_\xi = \infty$, the first term is not finite. The numerator is increasing linearly in σ_ξ. As a reminder:

$$\hat{P} = \frac{\sqrt{4\sigma_\xi^2 \sigma_\nu^2 + (\sigma_\nu^2)^2} - \sigma_\nu^2}{2} \tag{46}$$

This expression is increasing in σ_ξ but at a rate less than linearly.

The second term is positive so the overall expression remains infinite when added to the first term. Therefore, at $\sigma_\xi^2 = \infty$, the term $\frac{\xi^*}{\sigma_\xi}$ is not finite and so the probability of a draw of $\frac{\xi}{\sigma_\xi}$ greater than this is zero. The probability of over-shifting is then zero.

A.5 Proposition 4

Consider an equilibrium in which the signal consumers received is X and both firms make the same decision to reveal or not reveal their costs. In both cases, the equilibrium is a mixed strategy Nash equilibrium in which firms are indifferent between all prices in the support of the equilibrium price distribution. Therefore, firm profits from revealing or not revealing will be equal to the profits from playing the price r in either case. Then, we can plug in the reservation price with certainty from Equation 6 into the profit function in Equation 3 to get profits from revealing:

$$\pi(r) = \frac{1-\mu}{2}((C + \frac{s}{A(\mu)}) - C) - d = \frac{(1-\mu)s}{2A(\mu)} - d \tag{47}$$

Next, we can plug in the reservation price from Equation 9 into this profit function to get profits from not revealing:

$$\pi(r) = \frac{1-\mu}{2}(X + E[\epsilon|p = r, \epsilon > \epsilon^*] + \frac{s}{A(\mu)} - C) = \frac{1-\mu}{2}(E[\epsilon|p = r, \epsilon < \epsilon^*] - \epsilon + \frac{s}{A(\mu)}) \tag{48}$$

Profits in Equation 48 are strictly decreasing in ϵ, while profits in Equation 47 are constant. Therefore, if there exists a value ϵ^* such that the firm is indifferent between revelation and not, then for any $\epsilon > \epsilon^*$ the firm strictly prefers revelation, and for any $\epsilon < \epsilon^*$, the firm strictly prefers no revelation. Setting these two profits equal, we can solve for that value:

$$\frac{(1-\mu)s}{2A(\mu)} - d = \frac{1-\mu}{2}(E[\epsilon|p = r, \epsilon < \epsilon^*] - \epsilon + \frac{s}{A(\mu)})$$

$$\epsilon^* = \frac{2d}{1-\mu} + E[\epsilon|p = r, \epsilon < \epsilon^*]$$

This implicitly defines ϵ^*. A solution to this equation does not exist when $d = 0$ as $\epsilon^* \neq E[\epsilon|p = r, \epsilon < \epsilon^*]$. However, for a small d, this value will exist. It will not be well defined for sufficiently large d. Specifically, it will not exist when even the largest draw of ϵ, $\bar{\epsilon}$ does not induce the firm to reveal. Specifically when:

$$\frac{(1-\mu)s}{2A(\mu)} - d < \frac{1-\mu}{2}(E[\epsilon|p=r] - \bar{\epsilon} + \frac{s}{A(\mu)})$$

$$d > \frac{1-\mu}{2}(\bar{\epsilon} - E[\epsilon|p=r])$$

Given Assumption 3, d lies in the interval in which ϵ^* is well defined.

For draws of ϵ such that $\epsilon \geq \epsilon^*$, the firm (weakly) prefers to reveal. In this case, the conditions for Proposition 1 are met and $\rho = 1$.

For draws of ϵ such that $\epsilon < \epsilon^*$, the firm prefers not to reveal. In this case, the average transacted price is given by Equation 10.

$$\bar{p} = X + \mu\epsilon + (1-\mu)(\hat{\epsilon} + \frac{s}{A(\mu)})$$

However, each of the terms must be recalculated given the consumers knowledge that the firm did not reveal, and thus $\epsilon < \epsilon^*$. Denote the consumer's expectation of ϵ conditional on it being less than ϵ^*, ϵ^{**} ($\epsilon^{**} = E[\epsilon|\epsilon < \epsilon^*] = \frac{\int_{\underline{\epsilon}}^{\epsilon^*} \epsilon h(\epsilon) d\epsilon}{F(\epsilon^*)}$). This term must be negative because the unconditional expectation of ϵ is 0. The firms expectation of cost is now $X + \epsilon^{**}$. The ϵ term is now $C - X - \epsilon^{**} = \epsilon - \epsilon^{**}$. The $\hat{\epsilon}$ term will change relative to when the consumer did not know that $\epsilon < \epsilon^*$ but it remains a constant with respect to the actual draw of ϵ. Therefore it does not affect pass-through. Then rewriting the average price:

$$\bar{p} = X + \mu\epsilon + (1-\mu)(\hat{\epsilon} - \epsilon^{**} + \frac{s}{A(\mu)})$$

Then, because pass-through is linear in X and ϵ and the average value of ϵ in this region is ϵ^{**} which we know to be negative, we know that the average pass-through rate in this region is less than one.

Lastly, we know that both firms will reveal their cost when $\epsilon \geq \epsilon^*$. Because the decision to reveal is only relevant when the firm is facing a customer without the other firm's price quote, the other firm's decision to reveal has no effect on the firm's incentives. Therefore, if one firm finds it optimal to (not) reveal, the other firm will also find it optimal to (not) reveal.

Therefore in equilibrium, for a given there exists an equilibrium such that for a given signal x, there is a value of actual cost $c^* \equiv x + \epsilon^*$ above which all firms reveal their costs to consumers and below which they do not. Average pass-through rate is equal to one when costs are above c^* and average pass-through rate is lower than one when costs are below c^*.

Online Appendix B Robustness

B.1 Alternative Notions of Price

There exist alternative methods of calculating the "price paid" by borrowers. The first and simplest of these conversions is to use the price of "discount points" typically offered to borrowers. When a borrower takes out a mortgage, the lender generally gives them the ability to "pay points" or to make an additional upfront payment in exchange for a lower interest rate. As an example, a borrower could be offered a 4.25% interest rate with no points paid, or they could pay an additional 1% of the face value of their loan upfront in exchange for a 4% interest rate. In this example, the borrower paid 1 point and received a 25 basis point discount on their rate. This exchange rate of points to interest rate can vary across lenders and time, but a 1 point to 20-25 basis point rate serves as a good rule of thumb[22]. This methodology is nice in that it represents surplus to the borrower in a way that is directly accessible to a real borrower. If costs increased such that offered interest rates increased by roughly 20 basis points, a borrower could simply pay an additional point upfront and receive the same rate as before. The loss to the borrower resulting from the cost increase is just that point paid upfront. One downside to this method is that it is linear. As costs change more dramatically, this approximation might become less accurate.

Another way to convert the interest rate to an upfront payment is to use the prices facing the lender. After originating a loan, the lender must choose the coupon rate of the MBS pool that the loan will go into. These MBS generally only trade in 50 basis points increments (e.g. 3.0%, 3.5%, etc.). As the interest rate charged by the lender increases,

[22]Fuster, Lo, and Willen (2017) finds an exchange rate of "around 20 basis points" while Agarwal, Ben-David, and Yao (2017) finds an average of 22 basis points. In my main specification, I use 22 basis points.

they can place these loans into higher coupon securities, which trade at higher prices. One way to convert the interest rate paid by the borrower into an upfront price is therefore to interpolate between the differences of these prices of MBS of each coupon. If costs increased in such a way that offered interest rates increased by 50 basis points, this would allow lenders to place the loans in a higher coupon MBS and earn the difference between the two prices. To quantify the harm or benefit to the borrower from a change in interest rates, I use the 3.5% TBA price as a basis (the median coupon in the data) and a 4.25% interest rate (roughly the minimum interest rate for a loan that can be placed in a 3.5% MBS). I normalize the benefit to the borrower of having a 4.25% interest rate as zero. I then multiply the difference in the interest rate by the difference in the price of the next highest or lowest coupon MBS (divided by 0.5). This method maintains the benefit of using real market prices and it also captures some of the non-linearity of the relationship between interest rates and upfront payments. The price difference between coupons decreases as the coupon rate increases, meaning the rate of conversion between rate and upfront payment for this measure decreases as interest rate increases. A downside to this method is that TBA prices do not reflect an actual payment to the borrower.

In the following section, I estimate pass-through using each of these different measures of price (and find qualitatively similar results).

B.2 Regression Specifications

In this section, we repeat much of the analysis from Section 5, varying some of the assumptions made in estimation. Throughout this section, we find that the results presented in Section 5 are robust to the assumptions tested.

Table A.1 runs the same 6 regressions as Table 1 in the main text, but using the borrowers' points and fees paid to get to an interest rate of 4.25 as the price. I add the points and fees that the borrower actually paid to 0.22 times the difference between the transacted interest rate and 4.25. This assumes that borrowers are able to transact upfront points for a discount on their interest rate at the average exchange rate found in Agarwal, Ben-David, and Yao (2017). This measure would represent the borrowers' upfront payment if they always took out the same interest rate and used discount points to achieve that rate.

The pass-through estimates using this notion of price are scaled up slightly relative to those in Table 1, but the relative comparisons are qualitatively similar. We can still reject the null hypothesis that $\rho_1 = \rho_2$ at a 1% level under every specification. Thus, we can be confident that the pass-through rate of TBA prices to the borrowers' price measured in points is different than the pass-through rate of specified pool payups to that price. We generally cannot reject the null hypothesis that $\rho_1 = \rho_3$, though we can do so for specification 6 at the 5% level. It does not seem likely that the pass-through of TBA prices and LLPAs is different.

Table A.2 repeats the analysis but using the difference in TBA prices across coupons to value the differences in interest rates. As discussed above, I take the median gap between TBA price at each coupon, and then calculate an upfront value of the difference between the borrowers interest rate and 4.25 by interpolating between these prices. I then add this to the points and fees paid by the borrower. This notion of price represents the price of the borrowers' loan as valued by the TBA market.

In this table, pass-through estimates are scaled down relative to those using the first two notions of price. Again, the null hypothesis that $\rho_1 = \rho_2$ can be rejected across all relevant specifications. The null hypothesis that $\rho_1 = \rho_3$ is rejected only in the specifications that utilize LTV variation, and the sign of the difference is the opposite of what is found using the other two notions of price.

Table A.3 tests the sensitivity of the results to the loan size window used for the regression discontinuity. Each column of Table A.3 estimates the same regression equation as Table 1, column 1, but varies the loan size window used for the regression discontinuity. In each column, we report the estimates for ρ_1 and ρ_2 (pass-through of TBA price and specified pool payup, respectively), using a different subset of the data. Column 1 reports estimates using loans that have an original loan size within \$5,000 of a specified pool cut-off. These are the results presented in Table 1. Columns 2-4 report estimates from using loan sizes within \$1,000, \$500, and \$100 of the cut-offs, respectively. The estimates from these specifications are increasingly noisy due to decreased power, though all conclusions are identical. Most notably we can still reject the hypothesis that $\rho_1 = \rho_2$ at the 1% level. Finally, to ease concerns about borrowers (or lenders) manipulating loan sizes to be directly below the cut-offs, we estimate the same equation using loans that are within \$5,000 of a cut-off but **not** within \$500 of the cut-off. Again, we can clearly reject the equality of the two pass-through rates.

Table A.4 reports the results of estimating pass-through separately for each specified pool cut-off. One could worry that borrowers of different loan sizes have different demand for mortgages and thus different pass-through. Aggregating the results across each of the specified pool cutoffs as we do in the main specification increases the power of the test but also increases the risk of misspecification. We therefore run the same regression as in Table 1, column 1, but using only the loans near a single cut-off at a time. This makes the assumption that borrowers in the treated group (those in the higher pay-up specified pool) are the same as those in the untreated group more credible. In Table A.4 we can see that the results are quantitatively similar across each cut-off. The lower sample size makes each individual estimate noisier, but in each case the point estimate for TBA price pass-through is far greater than the estimate for specified pool payup pass-through. We can reject the null hypothesis that they are equal for LLB, MLB, and 200k eligible loans, but not for HLB or HHLB. For those two categories of loans, the standard errors on the estimate are quite wide due to low power. Taking all of these results together, it appears that the main result of different component-specific pass-through is not an artifact of aggregating across specified pool cut-offs. The result holds across each individual cut-off, though it is noisier and is not conclusive for two of the cut-offs when the results are viewed in a vacuum.

Last, Table A.5 reports the results of estimating each of the specifications in Table 1, but using a different MBS coupon rate as the basis for analysis. To estimate pass-through of TBA prices and payups to the prices paid by borrowers, a researcher needs to take a stance on which coupon MBS the loan is going into. In most of the analysis, I use the 3.5% coupon because this is the median coupon during the period studied. However, Table A.5 demonstrates that using the 4% coupon for TBA prices and payups does not change the main conclusion. We can still reject the hypothesis that the pass-through rates of TBA prices and payups under all specifications. Here, the pass-through of LLPAs is relatively lower than when 3.5% coupon MBS are used, although this varies across specifications and this hypothesis cannot be rejected across each of them. Overall, it is clear that the hypothesis that the pass-through of TBA prices and payups are different is not caused by the choice of base coupon.

	(1) Price	(2) Price	(3) Price	(4) Price	(5) Price	(6) Price
TBA Price	1.05	1.20	1.22	1.05	1.05	1.17
	[0.91,1.20]	[1.05,1.36]	[1.07,1.37]	[0.92,1.17]	[0.91,1.20]	[0.99,1.35]
Payup	0.44			0.45	0.45	0.25
	[0.16,0.73]			[0.17,0.72]	[0.17,0.72]	[0.035,0.47]
LLPA		1.03	1.14	1.03	1.06	0.95
		[0.91,1.15]	[1.03,1.25]	[0.72,1.34]	[0.93,1.20]	[0.84,1.07]
Loan Size Controls		✓	✓			
Credit Controls	✓					
LTV Restriction	≤ 80	$= 80$	≤ 80	$= 80$	≤ 80	≤ 80
Sample Window	Loan Size	FICO	FICO or LTV	Loan Size and FICO	Loan Size and FICO or LTV	All
N	69252	93344	216596	13636	37462	399260
R^2	0.685	0.689	0.695	0.685	0.686	0.704
P(TBA = Spec)	.003			.001	.002	0
P(TBA = LLPA)		.069	.366	.891	.884	.021

A.1: Transacted borrowers' price measured in discount points regressed on TBA prices, specified pool payups, and LLPAs. All regressions control for county fixed effects. All standard errors are clustered at the location, firm, time level. Point estimates for pass-through rates reported with 95% confidence intervals in brackets below. Loan size controls are fixed effects for loan size calculated using $50,000 buckets. Credit controls are fixed effects by LLPA bins. LTV restriction refers to whether the sample was restricted to borrowers with an LTV of 80 or not. Sample window refers to restriction of the sample to loans that are close to specified pool or LLPA cutoffs. Loan size refers to within $5000 of a specified pool cutoff, FICO refers to within 5 of an LLPA credit score cutoff, and LTV refers to within 2 of an LLPA LTV cutoff. N reports the sample size. P(TBA = Spec) reports the p-value of an F-test comparing the coefficients (pass-through rates) on TBA prices and specified pool payups. P(TBA = LLPA) reports the p-value of an F-test comparing the coefficients on TBA prices and LLPAs.

	(1) Price	(2) Price	(3) Price	(4) Price	(5) Price	(6) Price
TBA Price	0.51	0.58	0.59	0.53	0.51	0.56
	[0.45,0.57]	[0.52,0.65]	[0.52,0.65]	[0.48,0.58]	[0.46,0.57]	[0.48,0.63]
Payup	0.20			0.16	0.19	0.14
	[0.041,0.36]			[0.020,0.31]	[0.052,0.34]	[0.033,0.25]
LLPA		0.68	0.72	0.66	0.67	0.64
		[0.60,0.76]	[0.66,0.78]	[0.43,0.89]	[0.59,0.76]	[0.54,0.74]
Loan Size Controls		✓	✓			
Credit Controls	✓					
LTV Restriction	≤ 80	$= 80$	≤ 80	$= 80$	≤ 80	≤ 80
Sample Window	Loan Size	FICO	FICO or LTV	Loan Size and FICO	Loan Size and FICO or LTV	All
N	69,252	93,344	216,596	13,636	37,462	399,260
R^2	.541	.579	.578	.563	.542	.593
P(TBA = Spec)	.003			0	.001	0
P(TBA = LLPA)		.057	.001	.225	.001	.18

A.2: Transacted borrowers' price measured using TBA coupon spread regressed on TBA prices, specified pool payups, and LLPAs. All regressions control for county fixed effects. All standard errors are clustered at the location, firm, time level. Point estimates for pass-through rates reported with 95% confidence intervals in brackets below. Loan size controls are fixed effects for loan size calculated using $50,000 buckets. Credit controls are fixed effects by LLPA bins. LTV restriction refers to whether the sample was restricted to borrowers with an LTV of 80 or not. Sample window refers to restriction of the sample to loans that are close to specified pool or LLPA cutoffs. Loan size refers to within $5000 of a specified pool cutoff, FICO refers to within 5 of an LLPA credit score cutoff, and LTV refers to within 2 of an LLPA LTV cutoff. N reports the sample size. P(TBA = Spec) reports the p-value of an F-test comparing the coefficients (pass-through rates) on TBA prices and specified pool payups. P(TBA = LLPA) reports the p-value of an F-test comparing the coefficients on TBA prices and LLPAs.

	(1) Present Value	(2) Present Value	(3) Present Value	(4) Present Value	(5) Present Value
TBA Price	0.85	0.84	0.85	0.86	0.85
	[0.74,0.96]	[0.73,0.94]	[0.75,0.96]	[0.76,0.97]	[0.73,0.97]
Payup	0.37	0.41	0.40	0.38	0.36
	[0.14,0.60]	[0.18,0.64]	[0.17,0.64]	[0.15,0.62]	[0.13,0.59]
Loan Size Window Radius	5000	1000	500	100	5000 - 500 gap
N	69,252	21,488	15,437	12,835	53,294
R^2	.652	.67	.674	.673	.65
P(TBA = Spec)	.003	.007	.005	.003	.003

A.3: Transacted borrower present values regressed on TBA prices and specified pool payups. All regressions control for county and LLPA bin fixed effects. All standard errors are clustered at the location, firm, time level. Point estimates for pass-through rates reported with 95% confidence intervals in brackets below. Loan size window radius refers to restriction of the sample to loans that are close to specified pool cutoffs. I report the size of the window in dollars. 5000 - 500 gap refers to including loans within a window \$5000 above or below each cutoff but excluding loans within \$500 of a cutoff. P(TBA = Spec) reports the p-value of an F-test comparing the coefficients (pass-through rates) on TBA prices and specified pool payups.

	(1) Present Value	(2) Present Value	(3) Present Value	(4) Present Value	(5) Present Value
TBA Price	0.73	0.86	0.77	0.84	0.93
	[0.65,0.81]	[0.75,0.97]	[0.65,0.89]	[0.72,0.96]	[0.81,1.04]
Payup	0.39	0.31	0.55	0.51	0.35
	[0.20,0.58]	[0.17,0.46]	[0.27,0.84]	[0.16,0.87]	[0.047,0.65]
Spec Pool Cutoff	LLB	MLB	HLB	HHLB	200k
N	8,224	11,146	14,644	15,864	17,214
R^2	.629	.636	.657	.663	.663
P(TBA = Spec)	.009	0	.248	.143	.002

A.4: Transacted borrower present values regressed on TBA prices and specified pool payups. Results reported from estimating pass-through using only loans with a loan size within $5,000 of the cutoff for one specified pool story. Specified pool story cutoffs are: LLB ($85,000), MLB ($110,000), HLB ($150,000), HHLB ($175,000), and 200k ($200,000). All regressions control for county and LLPA bin fixed effects. All standard errors are clustered at the location, firm, time level. Point estimates for pass-through rates reported with 95% confidence intervals in brackets below. P(TBA = Spec) reports the p-value of an F-test comparing the coefficients (pass-through rates) on TBA prices and specified pool payups.

	(1) Present Value	(2) Present Value	(3) Present Value	(4) Present Value	(5) Present Value	(6) Present Value
TBA Price	1.19	1.41	1.42	1.21	1.19	1.36
	[0.91,1.48]	[1.14,1.67]	[1.15,1.69]	[0.95,1.47]	[0.90,1.48]	[1.00,1.71]
Payup	0.40			0.37	0.40	0.22
	[0.16,0.64]			[0.14,0.61]	[0.16,0.64]	[0.018,0.41]
LLPA		0.88	0.97	0.86	0.90	0.82
		[0.77,0.98]	[0.87,1.06]	[0.57,1.14]	[0.78,1.02]	[0.72,0.93]
Loan Size Controls		✓	✓			
Credit Controls	✓					
LTV Restriction	≤ 80	$= 80$	≤ 80	$= 80$	≤ 80	≤ 80
Sample Window	Loan Size	FICO	FICO or LTV	Loan Size and FICO	Loan Size and FICO or LTV	All
N	69252	93344	216596	13636	37462	399260
R^2	0.631	0.639	0.640	0.640	0.631	0.653
P(TBA = Spec)	.003			.001	.003	0
P(TBA = LLPA)		0	.002	.058	.046	.005

A.5: Transacted borrower present values regressed on TBA prices (where 4% coupon TBA contracts are used), specified pool payups, and LLPAs. All regressions control for county fixed effects. All standard errors are clustered at the location, firm, time level. Point estimates for pass-through rates reported with 95% confidence intervals in brackets below. Loan size controls are fixed effects for loan size calculated using $50,000 buckets. Credit controls are fixed effects by LLPA bins. LTV restriction refers to whether the sample was restricted to borrowers with an LTV of 80 or not. Sample window refers to restriction of the sample to loans that are close to specified pool or LLPA cutoffs. Loan size refers to within $5000 of a specified pool cutoff, FICO refers to within 5 of an LLPA credit score cutoff, and LTV refers to within 2 of an LLPA LTV cutoff. N reports the sample size. P(TBA = Spec) reports the p-value of an F-test comparing the coefficients (pass-through rates) on TBA prices and specified pool payups. P(TBA = LLPA) reports the p-value of an F-test comparing the coefficients on TBA prices and LLPAs.