

Measuring Information Frictions: Evidence from Capital Markets*

Andres Drenik

Columbia University

Juan Herreño

Columbia University

Pablo Ottonello

University of Michigan and NBER

October 29, 2019

Abstract

We study the importance of information frictions in asset markets. We develop a methodology to identify the extent of information frictions based on a broad class of models of trade in asset markets, which predict that these frictions affect the relationship between listed prices and selling probabilities. We apply our methodology to physical capital markets data, using a unique dataset on a panel of nonresidential structures listed for trade. We show that the patterns of prices and duration are consistent with the presence of asymmetric information. On the one hand, capital units that are more expensive because of their observable characteristics tend to have lower duration, as predicted by models of trading under a full information model. On the other hand, capital units that are expensive beyond their observable characteristics tend to have a longer duration, as predicted by models of trading under asymmetric information. Combining model and data, we estimate that asymmetric information can explain 21% of the +30% dispersion in price differences of units with similar observed characteristics. We quantify the effects of information frictions on allocations, prices, and liquidity, and show that the estimated degree of information frictions can lead to 15% lower output due to low trading probabilities of high-quality capital.

Keywords: Asymmetric information, asset markets, trading frictions, investment, fire sales, physical capital, misallocation.

*Drenik: ad3376@columbia.edu. Herreño: jdh2181@columbia.edu. Ottonello: ottonellopablo@gmail.com. We thank Mark Bils, Olivier Darmouni, Ricardo Lagos, Emi Nakamura, Stijn Van Nieuwerburgh, Guido Menzio, Venky Venkateswaran, Pei-Cheng Yu, Yu Zhu, conference participants at the Madison Money Meeting at the Wisconsin School of Business; Search and Matching Conference at Oslo; European Midwest Micro/Macro Conference at Chicago; Real Estate Academics conference at Columbia University; and seminar participants at Columbia University and The City University of New York, for very useful comments. Canyon Bosler and Nadim Elayan Balague provided excellent research assistance.

1 Introduction

A long tradition in economics and finance examines how asymmetric information affects asset markets. As information frictions affect the valuation, liquidity, and trading of assets, the last decade witnessed a surge of theories that identify these frictions as playing a central role in macroeconomic dynamics.¹ Motivated by these theories, our goal in this paper is to measure how relevant information frictions are to asset markets.

We develop a methodology to identify the extent of information frictions in asset markets and apply it to data from physical capital markets. Our key observation is that in a broad class of models of trade in asset markets, information frictions affect the relationship between listed prices and selling probabilities. On the one hand, under full information, high-quality assets attract more buyers than low-quality assets, which translates into higher prices and trading probabilities. On the other hand, under asymmetric information, sellers of high-quality assets signal their type and separate from sellers of low-quality assets. This implies that high-quality assets trade at high prices but low selling probabilities. We apply our methodology to the capital markets, using a unique dataset on a panel of nonresidential structures listed online for trade. We show that the patterns of prices and duration are consistent with the presence of asymmetric information. First, capital units that are more expensive because of their observable characteristics (e.g., location, size) tend to have lower duration, as predicted by models of trading under a full information model. Second, capital units that are expensive beyond their observable characteristics tend to have a longer duration, as predicted by models of trading under asymmetric information. Combining model and data, we estimate that asymmetric information can explain 21% of the +30% dispersion in the price differences of units with similar observed characteristics. We quantify the effects of information frictions on allocations, prices, and liquidity, and show that the estimated degree of information frictions can lead to 15% lower output due to low trading probabilities of high-quality capital.

The paper begins by laying out a model of trading in asset markets subject to information frictions that allows us to illustrate our methodology. The environment is aimed at capturing key aspects of our data, and focuses on a physical capital market in which sellers post prices and buyers search among listed units to produce. Sellers are agents who hold units of capital but cannot currently produce. Buyers are agents who can produce consumption units using capital as input. Gains from trade arise from these different production opportunities. Capital units are heterogeneous in their quality, i.e., in terms of the flow of consumption units they generate in production. Sellers choose a price at which to list and sell their assets in a decentralized market and

¹See, for example, [Guerrieri and Shimer \(2014\)](#), [House and Leahy \(2004\)](#), [Hendel and Lizzeri \(1999\)](#), [Eisfeldt \(2004\)](#), [Kurlat \(2013\)](#), and [Bigio \(2015\)](#)

face an endogenous probability determined by the relative mass of buyers and sellers. We show that the relationship between posted prices and probability of selling critically depends on the degree of asymmetric information on capital quality. When capital quality is public information, high-quality capital attracts more buyers and has a higher selling probability than low-quality capital. When capital quality is the private information of sellers, high-quality capital sellers choose to signal their type and separate from sellers of low-quality assets. They do so by choosing to list high-quality capital at such high prices that low-quality assets would not choose to mimic their pricing behavior; higher prices attract fewer buyers and are associated with lower trading probabilities. However, insofar as sellers can receive business opportunities in the future and start producing with their capital, high-quality sellers have a lower marginal cost of not trading than low-quality sellers. This separating equilibrium under asymmetric information resembles that of the classical model of [Spence \(1973\)](#), in which low types have a high marginal cost of effort and choose not to mimic the education levels of high types. In asset markets, the equivalent marginal effort exerted by high types is selling with a lower probability.

Based on this framework, we show that the extent of asymmetric information in asset markets can be identified from microlevel data on listed prices and time-to-sell of individual assets in a narrowly defined market.² The first important data elements are listed prices. If a researcher observes price differences between assets listed for trade within a narrowly defined market, those differences could reflect heterogeneous quality that is the private information of sellers. However, it could also reflect heterogeneity in asset quality that is known by all market participants, but unobserved by the researcher. Therefore, a second important element is the covariance between listed prices and duration. Under the hypothesis that there is no asymmetric information between these assets, one should observe a positive relationship between listed prices and time to sell in the data. This means that if one instead observes that higher prices tend to have a larger duration while listed, one can reject the hypothesis of full information in our framework. The more price differences are driven by differences in quality that are the private information of sellers, the higher the covariance between listed prices and duration should be. These different predictions mean that information frictions can be identified from the relationship between prices and selling probability across assets within a narrowly defined market.

We apply our methodology to data on physical capital markets. For this, we use a dataset that allows us to construct a novel joint measurement of individual market prices and duration of capital units listed for trade. In particular, our dataset contains the history of nonresidential structures

²In our model, assets within a market are perfect substitutes for the buyer. Therefore, by a narrowly defined market in the data we mean a group of listed assets that exhibit a high degree of substitutability for the buyer.

(retail and office space) listed for rent and sale in Spain by one of Europe’s main online real estate platforms, [Idealista](#), with rich information on each unit, including the listed price, exact location, size, age, and other characteristics. Given the data’s panel structure, for each unit we can compute the duration on the platform and the search intensity received, measured by the number of clicks received in a given month.

We begin the empirical analysis by providing a test of the model’s predictions. To this end, we first isolate the component of a property’s price that reflects the characteristics that are public information from the component that reflects the characteristics not observed by a researcher using these data (henceforth “the econometrician”), and potentially also by buyers. We estimate a hedonic regression of (log) prices per square foot on the set of characteristics included in each listing within a narrowly defined market (e.g., a neighborhood and compute the predicted prices from the hedonic regression and residual prices. We then study the comovement between predicted and residual prices with duration on the market. The data show (i) a negative relationship between predicted prices and duration and (ii) a positive relationship between residual prices and duration. The first empirical fact validates the prediction of the model under full information: Since predicted prices are obtained from observable characteristics, the theory predicts that on average, properties with better characteristics (which are reflected by a higher predicted price) should have a shorter duration on the market. The second fact provides evidence rejecting the hypothesis that there is no asymmetric information. That is, if residual prices would only reflect characteristics that are observed by market participants but not by the econometrician (e.g., listings that include pictures of the property), one would expect to observe a negative relationship between residual prices and duration, as observed for predicted prices and duration. We instead observe a positive relationship between residual prices and duration, which is consistent with the theory’s prediction about capital quality under asymmetric information.

We also provide an additional set of empirical results showing that our findings would be hard to rationalize with other theories of trading in asset markets that do not explicitly incorporate information frictions. First, we examine whether theories of price dispersion in markets with search frictions (e.g., [Burdett and Judd, 1983](#)) can rationalize our empirical findings. In principle, these theories generate a positive relationship between residual prices and duration: Sellers of homogeneous properties are indifferent between selling quickly at a low price or waiting in order to sell at a higher price, and thus randomize their choices. However, given the quantitative relation between residual prices and duration in the data, we show that any seller facing such price-duration trade-off will maximize the expected discounted revenues by choosing the highest price we see in the data. Second, we analyze whether heterogeneous sellers’ preferences can rationalize the empirical fact. To

do this, we repeat the analysis by computing the expected discounted revenues for a very broad set of preferences (discount factors from 0 to 0.99, and attitude toward risk from risk neutral to extreme forms of risk aversion). We find that all types of sellers would maximize their expected net present value by choosing the highest price observed in the data. Finally, we explore whether heterogeneous holding costs that sellers must pay can explain this fact, and conclude that in order for differential holding costs to explain the differences in expected discounted revenues in the data, they must be extremely large. It is worth mentioning that in addition, none of these alternative theories can rationalize the negative relationship between predicted prices and duration simultaneously with a positive relationship between residual prices and duration.

In the last part of the paper, we map the model to the data in order to quantify the extent of asymmetric information in the market for physical capital. Based on our identification strategy, we combine data on the standard deviation of residual prices and their covariance with duration to disentangle how much of the dispersion of residual prices reflects the characteristics of properties that are only known by the seller. Our calibration exercise shows that 21% of residual prices can be attributed to heterogeneous quality that is private information of the seller. To quantify the effects of asymmetric information, we compare the estimated model's predictions with the prediction of a model in which there is no private information. Asymmetric information has large effects on capital unemployment: The average unemployment rate of capital is 18% higher with asymmetric information. This is the result of asymmetric information reducing average trading probabilities, which increases the average unemployment rate. There is an additional effect that reduces aggregate utilization of capital: Asymmetric information reduces more the trading probabilities of high-quality capital, which increases the average quality of the pool of unemployed capital. This large effect on trading probabilities is translated into entrepreneurs' valuations due to an illiquidity discount. Therefore, the unconditional average price of a unit of capital is 16.7% lower due to asymmetric information. The overall welfare effect of information frictions is equivalent to an output loss of 18.4%. The magnitudes of the effects on allocations and prices are on the same order of magnitude of the estimated effects of search frictions (e.g., [Gavazza, 2016](#)).

Related Literature Our paper contributes to five branches of the literature. First is the literature on asymmetric information in asset markets. Classic theories show how information frictions can affect the quality of assets traded ([Akerlof, 1970](#)) and the financing investment opportunities (see, for example, [Stiglitz and Weiss, 1981](#); [Myers and Majluf, 1984](#)). We contribute to this literature by measuring these frictions in asset markets. We do so, by developing a methodology that builds on the theories of the asymmetric information in the presence of trading frictions pioneered

by Guerrieri et al. (2010) and further studied by Delacroix and Shi (2013), Guerrieri and Shimer (2014), and Lester et al. (2018), among others. We show that the effect on allocations of asymmetric information identified by these theories on allocations is large and relevant from a policy perspective.

Our second contribution is to the literature that measures asymmetric information. Important contributions to this literature include work by Chiappori and Salanie (2000), Ivashina (2009), and Einav et al. (2010), who measure asymmetric information in insurance and financial markets, and work by Kurlat and Stroebel (2015), who measure asymmetric information in housing markets. Our paper complements these studies, by developing a methodology to measure asymmetric information that exploits the relationship between prices and trading probabilities –which typically characterize asset markets– but that can be applied more broadly to other frictional markets.

Our third contribution is to the literature in trading frictions in asset markets. This includes the large body of work on financial markets (for a recent survey, see Lagos et al., 2017); housing markets (see, for example, Wheaton, 1990; Krainer, 2001; Caplin and Leahy, 2011; Piazzesi et al., 2015); and physical capital markets (see, for example, Kurmann and Petrosky-Nadeau, 2007; Gavazza, 2011; Cao and Shi, 2017; Ottanello, 2017). Our paper contributes to this literature by showing the relevance of the interaction between asymmetric information and trading frictions. In particular, asymmetric information can increase the duration of capital reallocation, leading to a potential scope of Pareto-improving market interventions.³

Finally, our paper contributes to the large body of research that measures resource misallocation (see, for example, Hsieh and Klenow, 2009). We contribute to this literature by studying a novel form of misallocation that stems from agents who own high-quality capital that signals their type by choosing to list capital at high prices, which are visited less frequently by buyers and have lower matching rates. This form of misallocation would typically not be measured in existing firms, but rather in unemployed capital.

Layout The rest of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 presents the data and empirical facts. Section 4 maps the model to the data and quantifies informational asymmetries. Section 5 presents countefactual analysis. Section 6 concludes.

³Empirical studies have shown that capital reallocation is large and procyclical (see, for example, Ramey and Shapiro, 1998, 2001; Eisfeldt and Rampini, 2006; Eisfeldt and Shi, 2018, and references therein). Related to these findings, recent work has studied the implications of secondary asset markets. See, for example, Lanteri (2018) for endogenous irreversibility and Gavazza and Lanteri (2018) for endogenous illiquidity.

2 Theoretical Framework

We construct a model of trading in asset markets subject to information frictions that allows us to illustrate our methodology. The environment is aimed at capturing key aspects of our data, and focuses on a physical capital market in which sellers post prices and buyers search among listed units to produce. We show that the relationship between posted prices and probability of selling critically depends on the degree of asymmetric information about capital quality, with capital of higher quality and price tending to match at higher rates under full information and at lower rates under asymmetric information.

2.1 Environment

Time is discrete and infinite, and there is no aggregate uncertainty.

Goods Two goods are traded: final goods and capital goods. Capital goods are heterogeneous in two dimensions: an observed quality $\omega \in \Omega \equiv [\omega_1, \dots, \omega_{N_\omega}]$, with $\omega_i < \omega$ for $i < j$, and an unobserved quality $a \in \mathcal{A} \equiv [a_1, \dots, a_{N_a}]$, with $a_i < a_j$ for $i < j$. As described in detail below, the unobserved quality a is the private information of the owner of the capital. Qualities in the economy are distributed jointly according to a probability distribution $\mathcal{G}(\omega, a)$, which is public information.

Markets Final goods are traded in a Walrasian market. Capital goods are traded in a decentralized market with search frictions. Sellers list their capital units in the decentralized market, with their quality ω perfectly observed by all market participants (below, we discuss the identification of the model when part of ω could be unobserved by the econometrician), and choose at what price q to post their units. Buyers dedicate labor to search and match, and can direct their search toward a submarket with a specific price q and a specific observed quality ω . The flow of new matches in submarket (ω, q) is given by $\mathcal{M}(k_t^s(\omega, q), h_t^s(\omega, q))$, where $k_t^s(\omega, q)$ and $h_t^s(\omega, q)$ denote, respectively, capital posted by sellers in submarket (ω, q) and period t and hours worked by buyers searching in submarket (ω, q) and period t . We assume that $\mathcal{M}(k^s, h^s) = \min\{\bar{m}(k^s)^\eta(h^s)^{1-\eta}, k^s\}$, where $\eta \in (0, 1)$ and $\bar{m} > 0$.⁴ In each submarket (ω, q) , the market tightness, denoted $\theta_t(\omega, q) \equiv \frac{h_t^s(\omega, q)}{k_t^s(\omega, q)}$, is defined as the ratio between buyers' hours of search and the mass of capital posted by sellers.⁵ Visiting submarket (ω, q) in period t , sellers face a probability $p(\theta_t(\omega, q)) \equiv \frac{\mathcal{M}(k_t^s(\omega, q), h_t^s(\omega, q))}{k_t^s(\omega, q)}$ of selling capital, and buyers match a mass of capital $\mu(\theta_t(\omega, q)) \equiv \frac{\mathcal{M}(k_t^s(\omega, q), h_t^s(\omega, q))}{h_t^s(\omega, q)}$ per hour of search.

⁴The assumed functional form of the matching function is convenient for tractability and is used frequently in related literature on labor search (e.g., Shimer, 2005).

⁵Following the directed search literature (see, for example Moen, 1997; Menzio and Shi, 2011), in submarkets that are not visited by any sellers, $\theta_t(\omega, q)$ is an out-of-equilibrium conjecture that helps determine equilibrium. See footnote 7 below for more details.

Agents, preferences, and technologies The economy is populated by a unit mass of entrepreneurs (also referred to as buyers) and a unit mass of capitalists (also referred to as sellers). Entrepreneurs have access to a technology to produce final goods using capital, given by $y_t(\omega, a) = \omega k_t(\omega, a)$, where $y_t(\omega, a)$ denotes its output in terms of final goods in period t and $k_t(\omega, a)$ denotes capital of type (ω, a) used as inputs for production in period t . Each period, entrepreneurs are also endowed with hours of work to dedicate to search activities in capital markets. They have preferences over consumption of the final good described by $\sum_{t=0}^{\infty} \beta^t \{c_{it} - \chi h_{it}\}$, where c_{it} denotes the consumption of agent i in period t and h_{it} the hours of work dedicated to search activities. Capitalists are agents who own capital goods but do not have access to a production technology, which gives rise to gains from trade between capitalists and entrepreneurs. They have preferences over consumption of the final good, described as $\sum_{t=0}^{\infty} \beta^t c_{it}$. Each period a capitalist faces a probability $\varphi_K \in (0, 1)$ of developing a business idea and becoming an entrepreneur. This assumption implies that capitalists with high-quality unobserved quality have a higher reservation value for these units, which is a frequent assumption in asymmetric information models of asset markets (e.g., [Akerlof, 1970](#)). To focus on a stationary equilibrium, and without loss of generality, we also assume that each period entrepreneurs have an i.i.d. probability $\varphi_E = \varphi_K$ of becoming capitalists, which can be seen as the exit rate of businesses.

Information and timing Capital quality a is the private information of the owner of the unit of capital. That is, buyers cannot distinguish capital units of different qualities a at a given price. The timing each period is as follows:

- i. Capitalists list prices for their capital units, which are perfectly observed by all agents.
- ii. Entrepreneurs purchase capital, choosing their search effort for capital units at different prices.
- iii. Entrepreneurs produce final goods using capital accumulated in current and previous periods.
- iv. Entrepreneurs exit and become capitalists with probability φ_K , and capitalists receive business ideas and become entrepreneurs with probability φ_E .

This setup requires that we specify the entrepreneur's beliefs about the type of capital, given a listed price. We assume that all entrepreneurs have the same beliefs. We describe beliefs by the mapping $\lambda_a(\omega, q) : \mathcal{A} \times \mathbb{R}_+ \rightarrow [0, 1]$, which denotes the probability that a unit of capital is of unobserved type a , given the price of capital q and the observed type ω . We also denote by $a^e(\omega, q) \equiv \sum_{a \in \mathcal{A}} a \lambda_a(\omega, q)$ the expected unobserved quality at the price q and observed quality ω , under beliefs $\lambda_a(\omega, q)$.

2.2 Optimization

Entrepreneurs An entrepreneur arrives to a period with capital holdings described by the matrix

$$\mathbf{k} \equiv \begin{bmatrix} k(\omega_1, a_1) & \dots & k(\omega_{N_\omega}, a_1) \\ \dots & \dots & \dots \\ k(\omega_1, a_{N_a}) & \dots & k(\omega_{N_\omega}, a_{N_a}) \end{bmatrix}, \text{ where } k(\omega, a) \text{ denotes capital with observed quality } \omega \in \Omega \text{ and}$$

unobserved quality $a \in \mathcal{A}$. The problem of an entrepreneur, in recursive form, is given by

$$V_{\mathcal{E}}(\mathbf{k}) = \max_{\{h(\omega, q) \geq 0\}_{\omega \in \Omega, q \in \mathbb{R}_+}} c - \chi \sum_{\omega \in \Omega} \int_{q \in \mathbb{R}_+} h(\omega, q) + \beta \hat{V}_{\mathcal{E}}(\mathbf{k}') \quad (\text{P1})$$

$$\text{s.t. } c + \sum_{\omega \in \Omega} \int_{q \in \mathbb{R}_+} q \mu(\theta(\omega, q)) h(\omega, q) dq = \sum_{\omega \in \Omega} \sum_{a \in \mathcal{A}} \omega a k(\omega, a)', \quad (1)$$

$$k(\omega, a)' = k(\omega, a) + \int_{q \in \mathbb{R}_+} \pi_a(\omega, q) \mu(\theta(\omega, q)) h(\omega, q) dq, \quad \forall (\omega \in \Omega, a \in \mathcal{A}), \quad (2)$$

where $k(\omega, a)'$ denotes capital of quality (ω, q) accumulated in the current period; $h(\omega, q)$ the hours worked searching for capital in submarket (ω, q) ; $\pi_a(\omega, q)$ the probability of matching a unit of capital of quality a when searching in submarket (ω, q) ; $V_{\mathcal{E}}(\mathbf{k})$ the expected lifetime utility for an entrepreneur from capital stocks \mathbf{k} ; $V_{\mathcal{K}}(\mathbf{k})$ the expected lifetime utility for a capitalist from capital stock \mathbf{k} (further described below); and $\hat{V}_{\mathcal{E}}(\mathbf{k}) \equiv (1 - \varphi_{\mathcal{E}})V_{\mathcal{E}}(\mathbf{k}) + \varphi_{\mathcal{E}}V_{\mathcal{K}}(\mathbf{k})$. Let $\tilde{h}(\omega, q; \mathbf{k})$ and $\tilde{q}(\omega, a; \mathbf{k})$ denote the policy functions associated with problem P1.

Capitalists The problem of a capitalist, in recursive form, is in turn given by

$$V_{\mathcal{K}}(\mathbf{k}) = \max_{s(\omega, a), q(\omega, a)} c + \beta \hat{V}_{\mathcal{K}}(\mathbf{k}') \quad (\text{P2})$$

$$\text{s.t. } c = \sum_{\omega \in \Omega} \sum_{a \in \mathcal{A}} s(\omega, a) p(\theta(\omega, q(\omega, a))) q(\omega, a) k(\omega, a), \quad (3)$$

$$k(\omega, a)' = [(1 - s(\omega, a)) + s(\omega, a)(1 - p(\theta(\omega, q(\omega, a))))] k(\omega, a), \quad \forall (\omega \in \Omega, a \in \mathcal{A}) \quad (4)$$

where $s(\omega, a) \in \{0, 1\}$ is an indicator variable that takes the value of one if the capitalist decides to participate in the decentralized market selling capital type (ω, a) ; $q(\omega, a)$ denotes the submarket choice of a capitalist of type (ω, a) ; and $\hat{V}_{\mathcal{K}}(\mathbf{k}) \equiv (1 - \varphi_{\mathcal{E}})V_{\mathcal{K}}(\mathbf{k}) + \varphi_{\mathcal{E}}V_{\mathcal{E}}(\mathbf{k})$.⁶ Equation (3) is the capitalist's budget constraint, which links consumption to the revenues from selling capital, given the choice of prices. Equation (4) is the capitalist's capital accumulation constraint. Capitalists in this setup must take into account that their choice of prices can signal their capital quality, which will be reflected in the equilibrium-tightness function $\theta(\omega, q)$. Let $\tilde{s}(\omega, a; \mathbf{k})$ and $\tilde{q}(\omega, a; \mathbf{k})$ denote the policy functions associated with problem P2.

⁶Allowing the capitalist to list in multiple submarkets for a capital unit of quality a would lead to the same equilibrium.

Linearity of recursive problems A tractable feature of the model is that the entrepreneur's and capitalist's recursive problems are linear in capital holdings, which follows from the linearity of preferences, production, capital accumulation, and search technologies. All proofs are in Appendix A.

Proposition 1. *Entrepreneurs' and capitalists' recursive problems are linear in capital stocks, i.e., value functions can be expressed as $V_{\mathcal{E}}(\mathbf{k}) = \sum_{\omega \in \Omega} \sum_{a \in \mathcal{A}} \nu_{\mathcal{E}}(\omega, a) k(\omega, a)$ and $V_{\mathcal{K}}(\mathbf{k}) = \sum_{\omega \in \Omega} \sum_{a \in \mathcal{A}} \nu_{\mathcal{K}}(\omega, a) k(\omega, a)$. The marginal value of capital holdings satisfies the recursive problems:*

$$\nu_{\mathcal{E}}(\omega, a) = \omega a + \beta \hat{\nu}_{\mathcal{E}}(\omega, a), \quad (\text{P3})$$

$$\nu_{\mathcal{K}}(\omega, a) = \max_{s(\omega, a), q(\omega, a)} s(\omega, a) p(\theta(\omega, q(\omega, a))) q(\omega, a) + s(\omega, a)(1 - p(\theta(\omega, q(\omega, a)))) \beta \hat{\nu}_{\mathcal{K}}(\omega, a), \quad (\text{P4})$$

where $\hat{\nu}_{\mathcal{E}}(\omega, a) \equiv (1 - \varphi_{\mathcal{K}}) \nu_{\mathcal{E}}(\omega, a) + \varphi_{\mathcal{K}} \nu_{\mathcal{K}}(\omega, a)$ and $\hat{\nu}_{\mathcal{K}}(\omega, a) \equiv (1 - \varphi_{\mathcal{E}}) \nu_{\mathcal{K}}(\omega, a) + \varphi_{\mathcal{E}} \nu_{\mathcal{E}}(\omega, a)$.

Proposition 1 implies that the value of a unit of capital of a given type does not depend on other capital holdings of its owner. For an entrepreneur, the value for a unit of capital of type (ω, a) is given by the utility flow generated by its production plus its continuation value, which takes into account the probability of exiting production. For a capitalist, the value of a unit of capital of type (ω, a) is given by the expected utility flow from selling the unit plus its continuation value, which takes into account the probability of receiving a business idea and becoming an entrepreneur. As a consequence of Proposition 1, policy functions do not depend on capital stocks, i.e., $\tilde{h}(\omega, q; \mathbf{k}) = h^*(\omega, q)$, $\tilde{s}(\omega, a; \mathbf{k}) = s^*(\omega, a)$, and $\tilde{q}(\omega, a; \mathbf{k}) = q^*(\omega, a)$.

Operating with recursive problems P3 and P4, one can solve for the marginal value of capital for entrepreneurs and capitalists:

Corollary 1. *The marginal value of capital is given by*

$$\nu_{\mathcal{K}}(\omega, a) = \frac{1}{\rho} \underbrace{[r_{\mathcal{K}}(\omega, a)(1 - \beta(1 - \varphi_{\mathcal{K}}))]}_{\substack{\text{expected revenue} \\ \text{from selling capital}}} + \underbrace{\beta \varphi_{\mathcal{E}} p_n(\omega, a) \omega a}_{\substack{\text{expected revenue from} \\ \text{future potential production}}}, \quad (5)$$

$$\nu_{\mathcal{E}}(\omega, a) = \frac{1}{\rho} \underbrace{[\omega a(1 - \beta(1 - \varphi_{\mathcal{K}})) p_n(\omega, a)]}_{\substack{\text{expected revenue} \\ \text{from production}}} + \underbrace{\beta \varphi_{\mathcal{E}} r_{\mathcal{K}}(\omega, a)}_{\substack{\text{expected revenue from} \\ \text{future potentially selling capital}}}, \quad (6)$$

where, $r_{\mathcal{K}}(\omega, a) \equiv s^*(\omega, a) p(\theta(\omega, q^*(\omega, a))) q^*(\omega, a)$ denotes the current revenue per unit of capital type (ω, a) for capitalists posting in an optimal submarket; $p_n(\omega, a) \equiv (1 - s^*(\omega, a)) + s^*(\omega, a)(1 - p(\theta(\omega, q^*(\omega, a))))$ the associated probability of not selling capital type (ω, a) ; and $\rho \equiv 1 - \beta(1 - \varphi_{\mathcal{K}}) +$

$p_n(\omega, a)(\beta^2(1 - \varphi_K - \varphi_E) - \beta(1 - \varphi_E))$ a discount rate that converts the period expected revenue into an expected lifetime value.

Corollary 1 shows that the expected discounted lifetime utility from a unit of capital for capitalists and entrepreneurs is both a function of the revenues from selling capital and the revenues from production. In the case of capitalists, their value from a unit of capital depends on the expected revenue of selling that unit of capital, scaled by the probability of continuing to be a capitalist, plus the revenue that capitalists have if they do not sell the unit of capital but instead develop a business idea, become entrepreneurs, and start producing with the unit of capital. In the case of entrepreneurs, their value from a unit of capital depends on the revenue from producing with that unit of capital, scaled by the probability of continuing to be an entrepreneur, plus the revenue from exiting the entrepreneurial activity and having to sell that unit of capital. Given that higher capital quality translates into higher production, this property of the value functions implies that capital quality strictly increases the value of capital for entrepreneurs, and weakly increases it for capitalists:

Corollary 2. *The value functions of entrepreneurs and capitalists are, respectively, increasing and non-decreasing in unobserved capital quality, i.e., $\frac{\partial v_E(\omega, a)}{\partial a} > 0$ and $\frac{\partial v_K(\omega, a)}{\partial a} \geq 0$.*

Free entry and equilibrium market-tightness function Proposition (1) characterizes the value of entrepreneurs' existing units of capital. In addition, Problem (P1) shows that entrepreneurs have free entry to purchase capital in all submarkets. Their optimality condition with respect to hours dedicated to search activities in all submarkets implies that

$$h^s(\omega, q)(\mu(\theta(\omega, q))(q - v_E^e(\omega, q)) + \chi)^+ = 0, \quad (7)$$

where $v_E^e(\omega, q) \equiv \sum_{a \in \mathcal{A}} \pi_a(\omega, q) v_E(\omega, a)$ denotes the expected value of purchasing capital at price q (which use the linearity result in Proposition 1). This condition requires that in open submarkets in which entrepreneurs are willing to search, the expected cost per unit of capital in that submarket $\left(q + \frac{\chi}{\mu(\theta(\omega, q))}\right)$ is equal to the expected value of the capital unit, $v_E^e(\omega, q)$.

For submarkets visited by a positive mass of capitalists, condition (7) determines the equilibrium market-tightness function $\theta(\omega, q)$. Following the directed search literature, we also assume that (7) also determines the market-tightness function in submarkets not visited by capitalists in equilibrium.⁷ This assumption implies that the equilibrium market-tightness function faced by

⁷ As argued by Menzio and Shi (2011), a justification of this assumption comes from considering a sequential game in which capitalists choose with a tremble their posted prices and then entrepreneurs choose in what submarkets to search. Because of the tremble, the tightness is well defined for all pairs (ω, q) . As we make the tremble arbitrarily small, $\theta(\omega, q)$ converges to (8).

capitalists is given by

$$\theta(\omega, q) = \mu^{-1} \left(\frac{\chi}{v_{\mathcal{E}}^e(\omega, q) - q} \right) \quad (8)$$

for all $q < v_{\mathcal{E}}^e(\omega, q)$, and $\theta(\omega, q) = 0$ for all $q \geq v_{\mathcal{E}}^e(\omega, q)$, meaning that capital units listed above the entrepreneurs' value of capital remain unmatched.

2.3 Equilibrium under Full Information

We first consider the case in which trading occurs with full information. This corresponds to the case with $\mathcal{A} = \{\bar{a}\}$. In this setup, with only one possible unobserved quality, entrepreneurs' beliefs are $\lambda_{\bar{a}}(\omega, q) = 1$. A competitive equilibrium in the economy under full information can then be defined as follows.

Definition 1 (Equilibrium under full information). An equilibrium under full information consists of value functions $v_{\mathcal{E}}^* : \Omega \rightarrow \mathbb{R}$ and $v_{\mathcal{K}}^* : \Omega \rightarrow \mathbb{R}$; policy functions $s^* : \Omega \rightarrow \{0, 1\}$ and $q^* : \Omega \rightarrow \mathbb{R}_+$; and a market-tightness function $\theta^* : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that:

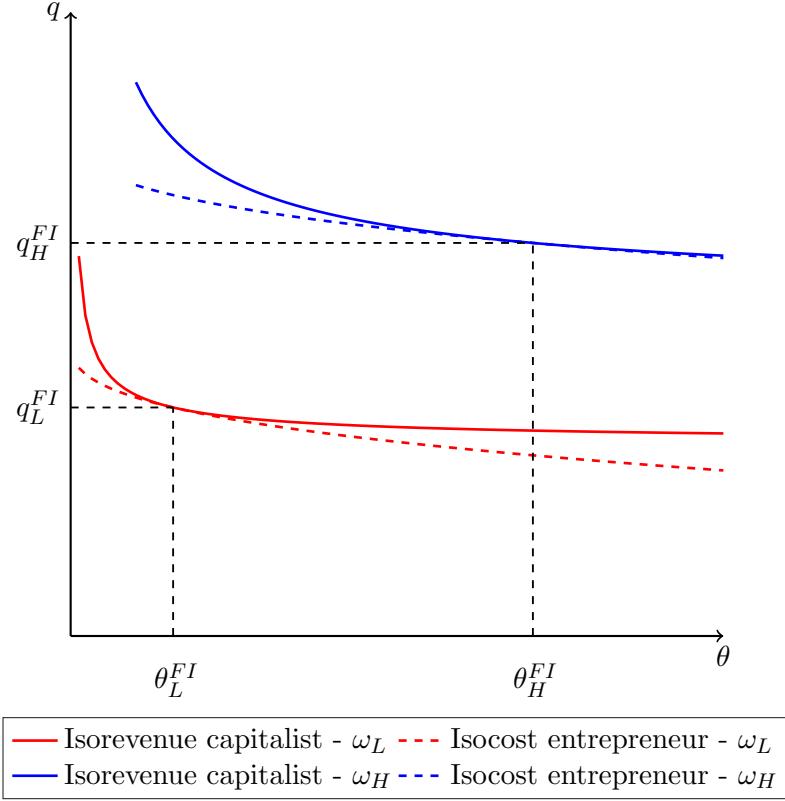
- i. $v_{\mathcal{E}}^*$ satisfies (P3) for all $\omega \in \Omega$ and $a = \bar{a}$.
- ii. $v_{\mathcal{K}}^*$ satisfies (P4) for all $\omega \in \Omega$ and $a = \bar{a}$; s^* and q^* are the associated policy functions.
- iii. The market-tightness function satisfies (8).

The following result derives the equilibrium price of capital and market tightness for each type of capital under full information.

Proposition 2. Under full information, there exists a unique equilibrium in which the price of capital and market tightness for capital quality ω are given by $q^{\text{FI}}(\omega, \bar{a}) = \eta v_{\mathcal{E}}(\omega, \bar{a}) + (1-\eta)\beta \hat{v}_{\mathcal{K}}(\omega, \bar{a})$ and $\theta^{\text{FI}}(\omega, \bar{a}) \equiv \theta(\omega, q^{\text{FI}}(\omega, \bar{a})) = (\bar{m}(1-\eta)\chi^{-1}(v_{\mathcal{E}}(\omega, \bar{a}) - \beta \hat{v}_{\mathcal{K}}(\omega, \bar{a})))^{\frac{1}{\eta}}$.

This result is graphically represented in Figure 1 for the case $\Omega = \{\omega_L, \omega_H\}$, with $\omega_L < \omega_H$. Dashed lines represent the isocost curves of entrepreneurs, with the one corresponding to the high quality ω_H simply being shifted upward by the difference in quality with ω_L . Solid lines denote the isorevenue curve of capitalists, with the one for the capitalist with high-quality capital having a lower slope. This is the result of requiring a lower "compensation" in terms of higher probability of sale for a given reduction in prices, since the outside option of capitalists is increasing in the quality of its capital. In equilibrium, capitalists choose the submarket that maximizes their utility subject to the entrepreneurs' indifference condition. Proposition 2 shows that under full information, the

FIGURE 1: Competitive Equilibrium under Full Information



price of a unit of capital and its matching rate are increasing in the quality of capital, which implies the following result.

Corollary 3. *In an equilibrium under full information, capital units with higher prices have higher matching rates. That is, if $\omega' > \omega$, then $q^{\text{FI}}(\omega', \bar{a}) > q^{\text{FI}}(\omega, \bar{a})$ and $p(\theta^{\text{FI}}(\omega', \bar{a})) > p(\theta^{\text{FI}}(\omega, \bar{a}))$.*

To see the intuition behind this result, replace the equilibrium price of capital in the equilibrium market-tightness function (8) to obtain

$$(1 - \eta)(\omega\bar{a} + \beta[\hat{\nu}_{\mathcal{E}}(\omega, \bar{a}) - \hat{\nu}_{\mathcal{K}}(\omega, \bar{a})]) = \frac{\chi}{\mu(\theta(\omega, q))}. \quad (9)$$

Equation (9) requires that in equilibrium, the net benefit for entrepreneurs from purchasing a unit of capital in the decentralized market relative to producing it must be equal to its search cost. As in standard models of directed search, the surplus (given by $\omega\bar{a} + \beta[\hat{\nu}_{\mathcal{E}}(\omega, \bar{a}) - \hat{\nu}_{\mathcal{K}}(\omega, \bar{a})]$) is “split” according to the elasticity of the matching function. Thus, since the price of capital scales with its productivity less than proportionally ($\eta < 1$), the net gain of purchasing capital is increasing in its productivity. By non-arbitrage, the search cost must be higher for capital units with higher productivity, meaning that sellers of these units match at a higher rate.

2.4 Equilibrium under Asymmetric Information

An equilibrium under asymmetric information is defined as follows:

Definition 2. A perfect Bayesian equilibrium consists of value functions $\nu_{\mathcal{E}}^* : \Omega \times \mathcal{A} \rightarrow \mathbb{R}$ and $\nu_{\mathcal{K}}^* : \Omega \times \mathcal{A} \rightarrow \mathbb{R}$; policy functions $s^* : \Omega \times \mathcal{A} \rightarrow \{0, 1\}$ and $q^* : \Omega \times \mathcal{A} \rightarrow \mathbb{R}_+$; a market-tightness function $\theta^* : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$; and a belief function $\lambda_a^* : \mathcal{A} \times \Omega \times \mathbb{R}_+ \rightarrow [0, 1]$, such that:

- i. $\nu_{\mathcal{E}}^*$ satisfies (P3) for all $(\omega, a) \in \Omega \times \mathcal{A}$.
- ii. $\nu_{\mathcal{K}}^*$ satisfies (P4) for all $(\omega, a) \in \Omega \times \mathcal{A}$, and s^* and q^* are the associated policy functions.
- iii. The market-tightness function satisfies (8), given beliefs $\lambda_a(\omega, q)$.
- iv. The belief function is derived from capitalists' strategies using Bayes' rule where possible.

Next, we consider the equilibrium under complete asymmetric information. This case corresponds to $\Omega = \{\bar{\omega}\}$ and $N_a > 2$. The last assumption is made in order to focus on the appropriate equilibrium selection mechanism (when $N_a = 2$ our selection mechanism still chooses the same type of equilibrium, but its requirements are stronger than needed). An important property of any equilibrium under asymmetric information is the following consequence of the single-crossing condition:

Lemma 1. In an equilibrium under asymmetric information, if $a > a'$ then $\theta(\bar{\omega}, q(\bar{\omega}, a)) \leq \theta(\bar{\omega}, q(\bar{\omega}, a'))$ for given $\bar{\omega}$.

This result is driven by the fact that the low-type capitalist has a higher marginal cost of not trading. A high type can never choose a higher probability of trading than the low type, because if the low type weakly prefers a low trading probability to a high trading probability, the high type would strictly prefer it.

In the setup considered, there are many equilibria, each supported by appropriate out-of-equilibrium beliefs. Of particular interest are separating equilibria, which satisfy the following additional condition.

Definition 3. A separating equilibrium under asymmetric information is an equilibrium in which different types post different prices: If $a \neq a'$ then $q(\bar{\omega}, a) \neq q(\bar{\omega}, a')$ for all pairs $(a, a') \in \mathcal{A}$.

An implication of this definition is that prices reveal the quality of each unit of capital. In a separating equilibrium, two conditions must be satisfied for any pair $(a, a') \in \mathcal{A}$ with $a > a'$. First,

it must be that the lower type a' does not want to mimic the higher type a ,

$$\begin{aligned} & p(\theta(\bar{\omega}, q(\bar{\omega}, a')))q(\bar{\omega}, a') + (1 - p(\theta(\bar{\omega}, q(\bar{\omega}, a'))))\beta\nu(\bar{\omega}, a') \\ & \geq p(\theta(\bar{\omega}, q(\bar{\omega}, a)))q(\bar{\omega}, a) + (1 - p(\theta(\bar{\omega}, q(\bar{\omega}, a))))\beta\nu(\bar{\omega}, a'), \end{aligned} \quad (10)$$

and that the higher type a does not want to mimic the lower type a' ,

$$\begin{aligned} & p(\theta(\bar{\omega}, q(\bar{\omega}, a)))q(\bar{\omega}, a) + (1 - p(\theta(\bar{\omega}, q(\bar{\omega}, a))))\beta\nu(\bar{\omega}, a) \\ & \geq p(\theta(\bar{\omega}, q(\bar{\omega}, a')))q(\bar{\omega}, a') + (1 - p(\theta(\bar{\omega}, q(\bar{\omega}, a'))))\beta\nu(\bar{\omega}, a). \end{aligned} \quad (11)$$

The definition of a separating equilibrium does not impose any constraints on off-equilibrium beliefs. For example, any equilibrium satisfying (10) and (11) could be supported with these off-equilibrium beliefs: After observing any off-equilibrium choice $q(\bar{\omega}, a)$, the entrepreneur believes that the deviation was made by the lowest quality a_1 . With such beliefs and the optimal response of the entrepreneur, no capitalist has an incentive to deviate. The following result imposes more structure on these beliefs by considering the equilibrium that satisfies the *D1 criterion* of Cho and Kreps (1987). What this criterion does is to first find the set of types that are more likely to deviate relative to the equilibrium choices. After requiring that entrepreneurs have beliefs consistent with this set after observing the deviation, the *D1 criterion* eliminates equilibria in which a capitalist's payoff of the deviation under the worst entrepreneurs' consistent belief is not equilibrium dominated.

Proposition 3. *There exists a unique equilibrium that satisfies the D1 criterion. This is a separating equilibrium, in which the tightness function $\theta^{\text{AI}}(\bar{\omega}, a) \equiv \theta(\bar{\omega}, q^{\text{AI}}(\bar{\omega}, a))$ and prices $q^{\text{AI}}(\bar{\omega}, a)$ satisfy:*

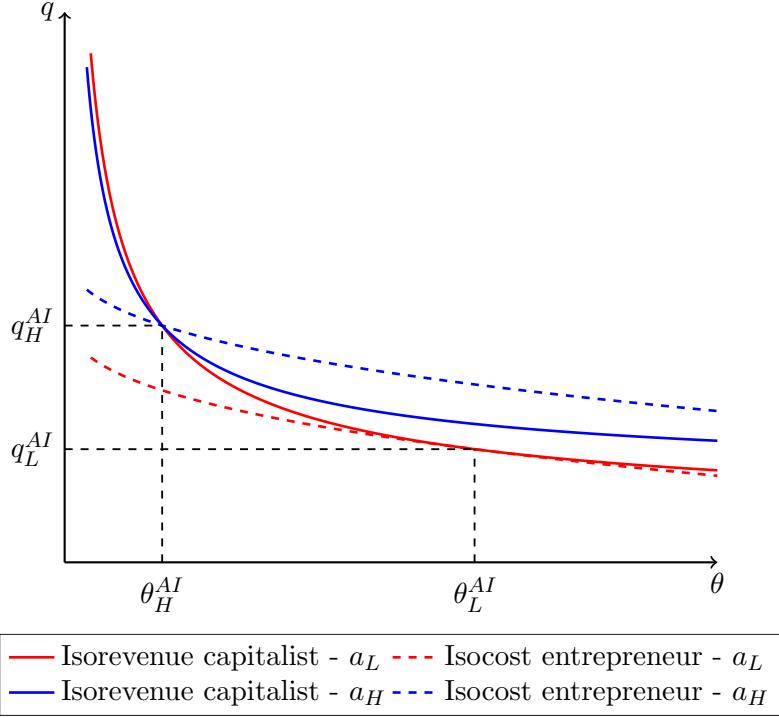
(i) *For $i = 1$,*

$$q^{\text{AI}}(\bar{\omega}, a_i) = \underset{q(\bar{\omega}, a_i)}{\text{argmax}} p(\theta(\bar{\omega}, q(\bar{\omega}, a_i)))q(\bar{\omega}, a_i) + (1 - p(\theta(\bar{\omega}, q(\bar{\omega}, a_i))))\beta\nu(\bar{\omega}, a_i).$$

(ii) *For all $a_i \in \mathcal{A}$ and $i \geq 1$,*

$$\begin{aligned} & p(\theta^{\text{AI}}(\bar{\omega}, a_i))q^{\text{AI}}(\bar{\omega}, a_i) + (1 - p(\theta^{\text{AI}}(\bar{\omega}, a_i)))\beta\nu(\bar{\omega}, a_i) \\ & = p(\theta^{\text{AI}}(\bar{\omega}, a_{i+1}))q^{\text{AI}}(\bar{\omega}, a_{i+1}) + (1 - p(\theta^{\text{AI}}(\bar{\omega}, a_{i+1})))\beta\nu(\bar{\omega}, a_i). \end{aligned}$$

FIGURE 2: Competitive Equilibrium under Asymmetric Information



(iii) For all $a_i \in \mathcal{A}$,

$$\theta^{\text{AI}}(\bar{\omega}, a_i) \equiv \theta(\bar{\omega}, q^{\text{AI}}(\bar{\omega}, a_i)) = \mu^{-1} \left(\frac{\chi}{\nu(\bar{\omega}, a_i) - q^{\text{AI}}(\bar{\omega}, a_i)} \right).$$

This result was originally established by Guerrieri et al. (2010) in a search model with adverse selection under a different equilibrium selection mechanism. We show that a similar equilibrium is obtained in a signaling model. The intuition behind this result can be seen in Figure 2 for the case $\mathcal{A} = \{a_L, a_H\}$, with $a_L < a_H$. In a separating equilibrium, the outcome in the submarket for the lowest quality capital is the same as the one obtained under full information (see Figure 1). However, the outcome in the submarket for higher-quality capital is distorted by the fact that higher-quality capitalists maximize expected utility subject to the constraint that lower-quality capitalists do not have a strict preference for participating in their submarket and given entrepreneurs' belief function. Although multiple separating equilibria are supported by different sets of beliefs (see the proof of Proposition 3 in the Appendix), the unique separating equilibrium that satisfies the *D1 criterion* is the one depicted in Figure 2. In such an equilibrium, the higher-quality capitalist chooses the allocation that renders the lower-quality capitalist just indifferent between mimicking and not, because it minimizes the signaling "effort" imposed by the lower trading probability. Thus, the following simple result follows from the separating equilibrium, with important implications that can be tested in the data.

Corollary 4. *In the unique separating equilibrium with asymmetric information that satisfies the D1 criterion, capital units with higher prices have lower matching rates: If $a > a'$, then $q^{\text{AI}}(\bar{\omega}, a) > q^{\text{AI}}(\bar{\omega}, a')$ and $p(\theta^{\text{AI}}(\bar{\omega}, a)) < p(\theta^{\text{AI}}(\bar{\omega}, a'))$.*

The model predicts that owners of capital units that possess characteristics that are not possible to announce in a listing signal their type by choosing a high price, which is associated with low trading probabilities and is not easy to mimic by owners of low-quality capital, given their higher marginal cost of not trading. For this reason, while the market tightness faced by the lowest type is the same under full and asymmetric information ($\theta^{\text{AI}}(\bar{\omega}, a') = \theta^{\text{FI}}(\bar{\omega}, a')$), the tightness faced by the higher type $a > a'$ is lower under asymmetric information than under full information ($\theta^{\text{AI}}(\bar{\omega}, a) < \theta^{\text{FI}}(\bar{\omega}, a)$). This implies an important qualitative change in the relationship between prices and duration on the market created by asymmetric information and signaling.

2.5 Identification

In the previous subsections, we have shown how prices and trading probabilities respond differently to changes in the observed and unobserved components of the quality of a unit of capital. Given the implications that asymmetric information has on the allocation, a natural question is: Can we use micro-data on prices and duration on the market to estimate the quantitative relevance of information frictions in the capital market? A priori, one could think that by using data on observables (e.g., location, size, number of rooms, etc.) as proxies for the observed component of the quality of capital ω , one could estimate the dispersion of the unobserved component a by looking at the dispersion of residual prices (i.e., the residual of a regression of prices on observables). However, suppose instead that a component of ω is in fact observed by market participants, but not “by the econometrician.” More formally, suppose that ω is a function of characteristics observed only by market participants z and a vector of characteristics observed by everyone x (including the econometrician). Then, this exercise will not provide the right answer. Here, we show that despite the potential presence of factors unobserved by the econometrician, we can still recover two key objects: the aggregate dispersion of the z and a components and their relative magnitudes. In other words, we develop a methodology to recover the importance of asymmetric information using micro-data on prices and duration.

Our theoretical identification analysis is conducted under the assumption that the elasticity of the matching function $\eta = 0.5$. In this special case, the model has a closed-form solution that allows us to show our identification result and facilitates exposition of the intuition. In the analysis of the extended model for the quantitative analysis, we provide a numerical exploration of the identification argument to show that it is valid there.

First, we present the assumptions we make in order to identify the presence of asymmetric information in the data:

Assumption 1. A unit of capital with observable characteristics x and unobserved components z and a produces $\exp(x'\delta)za$ efficiency units of capital.

Assumption 2. $(z, a) \perp\!\!\!\perp x$.

Assumption 3.

$$\begin{pmatrix} \log z \\ \log a \end{pmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_z^2 & \sigma_{z,a} \\ \sigma_{z,a} & \sigma_a^2 \end{bmatrix} \right).$$

The first assumption imposes restrictions on how the quality of a unit of capital gets transformed into efficiency units.⁸ The second assumption requires that the components of quality that are unobserved by the econometrician are independent of the observable characteristics x . However, as will become clear below, this assumption is not necessary for identification. In fact, one could allow for correlation between z and x , and a and x and still identify the parameters of interest with the use of additional moments from the data. However, by disposing from this assumption, one would need to impose structure on such correlations. In the spirit of tractability, we abstract from this possibility. The last assumption is more operational, and assumes that the unobserved heterogeneity is jointly log-normally distributed.⁹

We assume that the other parameters from the model are calibrated using aggregate data, including the discount factor β , the parameters of the matching function \bar{m} and η , the search cost χ , the outside option of the seller ϕ , and the decreasing returns in the production function α . Thus, the goal is to separately identify the parameters of the covariance matrix (σ_ω^2 , σ_a^2 , and $\sigma_{\omega,a}$) and the vector δ affecting the transformation into efficiency units of capital, using the micro-data.

The identification argument proceeds by first showing that ratios of prices and (expected) duration across units of capital only depend on ratios of different x 's, z 's, and a 's.

Proposition 4. If $\eta = 0.5$, then for any two units of capital (\bar{x}, \bar{z}, a_i) and $(\tilde{x}, \tilde{z}, a_{i+1})$, the unique equilibrium that satisfies the D1 criterion features prices, and ratios of prices and expected duration given by

$$q(\bar{x}, \bar{z}, a_i) = \exp(\bar{x}'\delta)\bar{z}a_i \left(1 - \frac{1}{2}\Theta_i \left(\frac{a_1}{a_2}, \dots, \frac{a_{i-1}}{a_i}, \phi, \beta \right) \right),$$

⁸This specific assumption is not needed. What is important is that x , z and a get transformed into efficiency units of capital in a *monotonic* way. That is, a certain feature of a unit of capital always increases or decreases the efficiency units of the unit.

⁹The means of both components are not separately identified from the average observable productivity. Thus, we include a constant term in the vector x and normalize the mean of both unobserved components to zero. This is without loss of generality, because the presence of asymmetric information is measured by the covariance matrix of z and a , and not by their means.

$$\frac{q(\bar{x}, \bar{z}, a_{i+1})}{q(\tilde{x}, \tilde{z}, a_i)} = \mathcal{Q} \left(\frac{\exp(\bar{x}'\delta)}{\exp(\tilde{x}'\delta)}, \frac{\bar{z}}{\tilde{z}}, \frac{a_{i+1}}{a_i}; \frac{a_i}{a_{i-1}}, \dots, \frac{a_2}{a_1}, \phi, \beta \right)$$

and

$$\frac{\mathbb{E}(D(\bar{x}, \bar{z}, a_{i+1}))}{\mathbb{E}(D(\tilde{x}, \tilde{z}, a_i))} = \frac{1/p(\theta(\bar{x}, \bar{z}, q(\bar{x}, \bar{z}, a_{i+1})))}{1/p(\theta(\tilde{x}, \tilde{z}, q(\tilde{x}, \tilde{z}, a_i)))} = \mathcal{D} \left(\frac{\exp(\bar{x}'\delta)}{\exp(\tilde{x}'\delta)}, \frac{\bar{z}}{\tilde{z}}, \frac{a_{i+1}}{a_i}; \frac{a_i}{a_{i-1}}, \dots, \frac{a_2}{a_1}, \phi, \beta \right),$$

with $\frac{\partial \mathcal{Q}(\cdot)}{\partial \bar{x}/\tilde{x}} > 0$, $\frac{\partial \mathcal{Q}(\cdot)}{\partial \bar{z}/\tilde{z}} > 0$, $\frac{\partial \mathcal{Q}(\cdot)}{\partial a_{i+1}/a_i} > 0$, $\frac{\partial \mathcal{D}(\cdot)}{\partial \bar{x}/\tilde{x}} < 0$, $\frac{\partial \mathcal{D}(\cdot)}{\partial \bar{z}/\tilde{z}} < 0$ and $\frac{\partial \mathcal{D}(\cdot)}{\partial a_{i+1}/a_i} > 0$.

Proposition 4 shows that while the vector δ is identified by the first moment of the distribution of prices, the remaining parameters are identified by the second moments of the joint distribution of prices and durations. First, notice that the vector δ can be obtained from a regression of (log) prices on observable characteristics. This is the result of Assumption 2. Second, given any two units of capital, the ratio of their prices and expected duration are only functions of the ratio of qualities x , z , and a . They also depend on parameters δ , ϕ , and β , but these are separately identified or calibrated (see Section 4 with the quantitative model). Third, the ratios of prices and duration depend on the z and x components of these two units only, while they depend on the a components of all units of lower quality. This is because prices fully reflect the components that are observed by market participants (x and z), while distortions created by asymmetric information accumulate over types. That is, asymmetric information distorts the relative allocation of capital units of quality a_1 and a_2 , which in turn affects the relative allocation of capital units of quality a_2 and a_3 , and so on.

More importantly, Proposition 4 shows that the ratio of both the z and the a components have similar effects on the ratio of prices, but opposing effects on the ratio of (expected) duration. This is the key identifying result. While larger gaps between the z and the a components of two units both increase price dispersion, a larger gap in the z component decreases relative duration, and a larger gap in the a component increases relative duration. These opposite relationships allow for identification of the dispersion and importance of the a component. While the joint dispersion of both components ($\sigma_z^2 + \sigma_a^2$) is disciplined by aggregate price dispersion, the relative dispersion of the z and a components (σ_z^2/σ_a^2) is disciplined by measures that summarize the comovement between a unit's price and its corresponding duration.

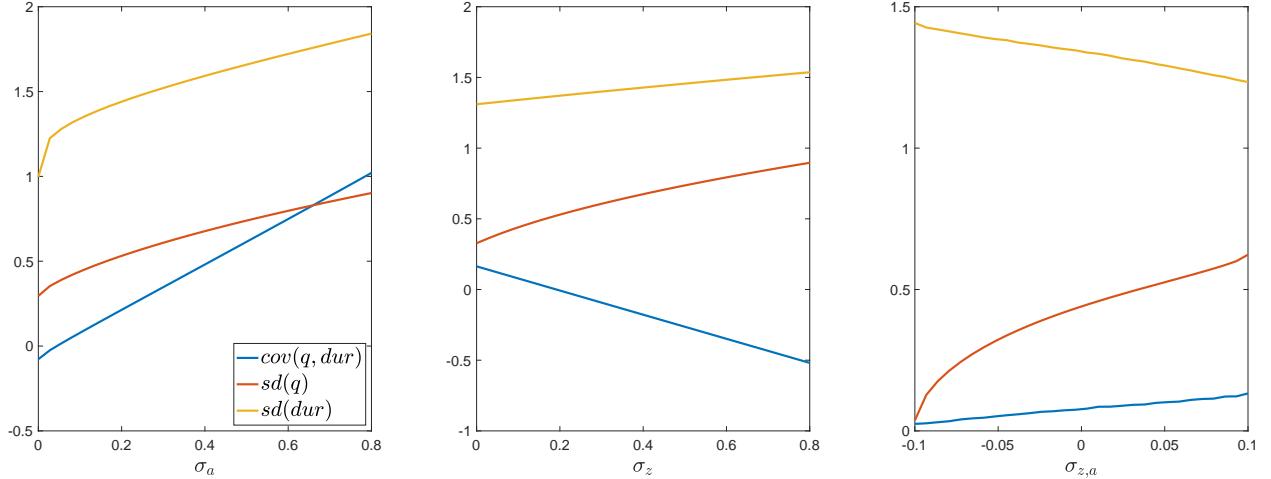
Finally, aggregate duration dispersion is informative of the covariance $\sigma_{z,a}$. While higher dispersion in both the z and a components of quality increase the dispersion of duration, a higher $\sigma_{z,a}$ decreases it. The intuition comes from the fact that duration is decreasing in z and increasing in a . Thus, to the extent that both components are positively correlated, their effects on expected duration are neutralized and the dispersion of duration should be small. On the other hand, if they

are negatively correlated, their effects would tend to go in the same direction and the dispersion of duration should be high.

In principle, there could be many data moments that are informative about the dispersion of prices, duration, and the comovement between them. We propose the following three standard moments: i) the standard deviation of (log) prices, ii) the standard deviation of (log) duration, and iii) the covariance between (log) prices and duration. Figure 3 plots the relationship between these three moments and the parameters of the log-normal distribution of (z, a) , and illustrates the intuition behind our methodology. Although these three moments are affected by the three parameters of interest, the figure shows how σ_a and σ_z can be identified using data on the covariance between prices and duration, and how the covariance $\sigma_{z,a}$ can be recovered using data on the standard deviation of duration on the market.

The figure also illustrates an important prediction of the model. As $\sigma_z \rightarrow \infty$ (or as $\sigma_a \rightarrow 0$), the covariance between prices and duration becomes negative, since the model behaves asymptotically as in the extreme full information case. On the other hand, as $\sigma_z \rightarrow 0$ (or as $\sigma_a \rightarrow \infty$), the covariance becomes positive, since the equilibrium behaves as in the full asymmetric information case. In the following empirical section, we exploit these results to devise a simple test for the presence of asymmetric information.

FIGURE 3: Illustration of Identification



3 Empirical Evidence

In this section, we first present and describe the micro data on commercial real estate from Spain. We then document the extent to which differences in capital prices can be explained by observable characteristics. Next, using micro-data on residual prices and measures related to search intensity, we provide suggestive evidence that rejects the null hypothesis of full information in the market of

physical capital. Finally, we present evidence showing that other potential mechanisms cannot be the main drivers behind the patterns observed in the data.

3.1 The Data

Our data consist of a panel of nonresidential structures (retail and office space) listed for sale and rent. The source of these data is [Idealista](#), one of Europe's leading online real estate intermediaries.¹⁰ The frequency of the panel is monthly, and it includes the universe of capital units that were listed on this platform between 2005 and 2018. The data contain information during the period of time each listing was active online. The dataset includes roughly 8.9 million observations for Spain, where an observation corresponds to a property–month pair. Overall, these observations come from over 1.15 million different capital units.

For each property, we observe a wide range of characteristics that we link to its price. In particular, we observe the address of the property, its construction year, its area, standardized self-reported condition of the unit, the number of rooms, and whether the property has heat or air conditioning. Table 1 presents some descriptive statistics on these characteristics.¹¹ The table includes four columns, which are the mean and standard deviation for properties listed for rent and sale, respectively. Our main variable of interest for each capital unit is its price, which we observe for each property at a monthly frequency. The average sale price per square foot is \$162 (expressed in constant 2017 dollars), and monthly rents are around \$1 per square foot per month. The properties are relatively old, with the average age around 26 years. The properties have similar sizes regardless of the operation.

For each capital unit, we also observe three key variables related to its time-to-sell and the attention it receives on the platform. First, based on the identifier of each property, we compute the number of months the unit is listed on the platform, which we refer to as *duration*. Table 1 shows that units for rent and sale remain on the platform 7.7 and 9.2 months on average, respectively. Second, we compute each capital unit's search volume in each month, measured by the number of views and clicks each listing received and the number of emails the seller receives from potential buyers through the platform. Each listing for rent and sale was on average viewed 1,460 and 875 times, respectively. Similarly, listings for rent and sale received 72 and 46 clicks per month, respectively, and 2.9 and 3 emails per month, respectively.

Appendix B provides more details on the data. In particular, Section B.1 describes how the

¹⁰[Idealista](#) is the leading online platform in the real estate market in Spain (see [Comparison of users](#) and [Comparison of platform](#)). Other papers in the literature have made use of data from online platforms in the real estate market (see, for example, ([Piazzesi et al., 2015](#))).

¹¹For those variables that change over time, we first take the average of the variable for each listing and report the average of that variable across listings.

TABLE 1: Descriptive Statistics

	Mean Rent	Std. Rent	Mean Sale	Std. Sale
Price	1.09	0.77	162.28	129.74
Duration	7.70	7.56	9.21	8.25
Construction Date	1986.61	19.67	1986.96	20.01
Area	2588.87	3977.47	2948.43	4507.27
New	0.00	0.06	0.04	0.21
Needs Restoration	0.08	0.28	0.15	0.36
Good Condition	0.91	0.28	0.81	0.39
Rooms	2.60	2.80	2.38	3.20
Restrooms	1.32	1.30	1.20	1.54
Heating	0.38	0.49	0.28	0.45
AC	0.77	0.42	0.71	0.46
Emails	2.90	2.09	2.99	2.05
Views	1460.25	2264.15	875.37	1333.88
Clicks	72.31	97.51	46.96	59.26
Number of Obs.	6.1e+05	6.1e+05	3.8e+05	3.8e+05

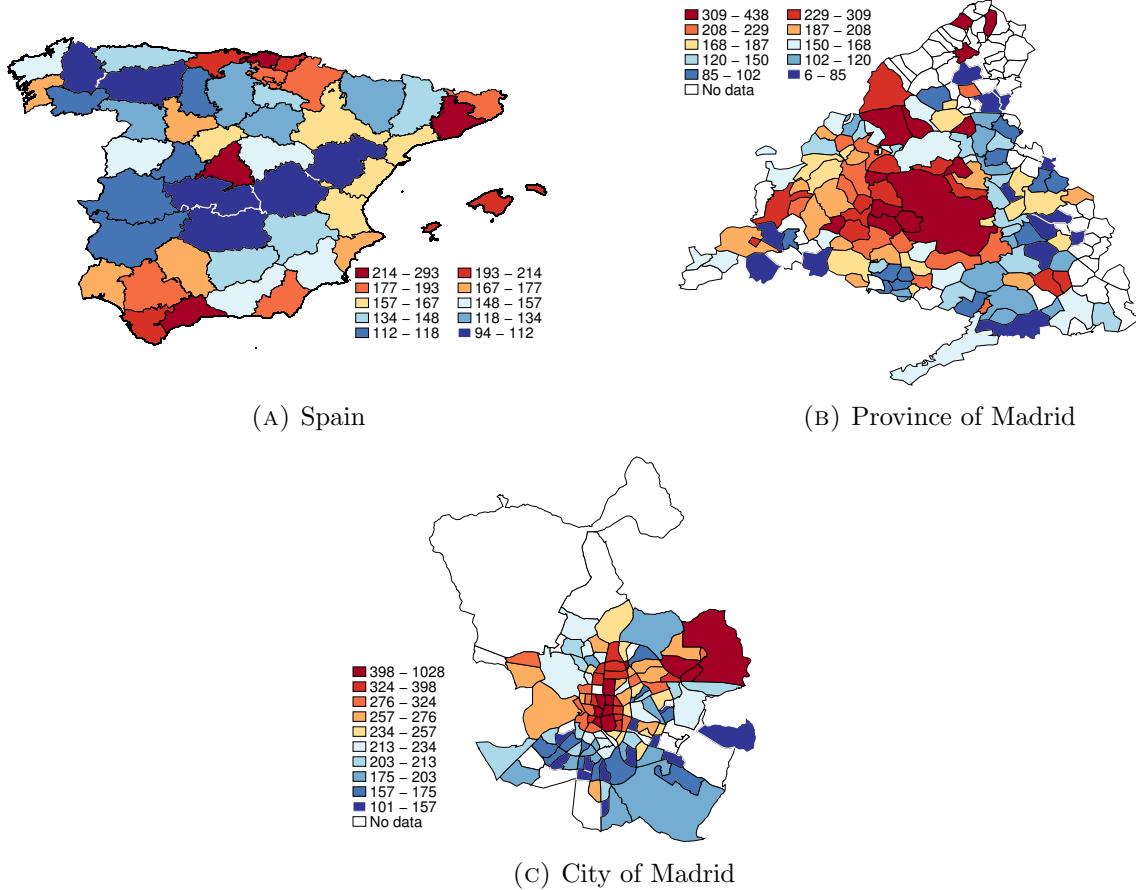
Note: Price is the price per square foot in constant 2017 dollars. Duration is the number of months a property lasted in the database. Construction date is the year in which the property was built. Property area is measured in square feet. “New” is a categorical variable that takes the value of 1 if the property is new. “Needs Restoration” is a categorical value that takes the value of 1 if the owner declares the property needs reparations. “Good Condition” takes the value of 1 if the property does not need reparations. “Rooms” is the number of separate rooms the property has, similar for “Restrooms”. Heating and AC are categorical variables that take the value of 1 when the property has some heating and air conditioning technologies. Emails is the number of times per month a property received an email from a potential customer. Views is the number of times a property appeared in the screen of a potential customer per month. Clicks is the number of times per month a potential customer clicked on the property listing to see its details.

online platform works. Section B.2 studies the representativeness of the dataset, showing that the data from the online platform are consistent with aggregate patterns observed in Spain during the period of analysis in terms of the aggregate evolution of prices and the timing of sales.

Variation over Time and Space. Here, we describe the coverage of the data and the observed variation in prices of listed capital. An important observable dimension of capital prices is the time dimension. Figure 5 shows that, as is well known, the price of capital units experienced large fluctuations over the last 15 years during the boom and bust in the real estate market in Spain. In the markets for both sale and rent, prices declined by more than 50% from the peak in 2007 to the trough in 2012. Since then, prices have remained stable. Another key observable dimension that explains differences in capital prices is location. Figure 4 shows remarkable differences in sale prices across regions at different levels of aggregation. Panel (a) shows the average price across the 50 provinces in Spain; Panel (b) zooms in on the province of Madrid and shows the average price across municipalities in that province; Panel (c) zooms in on the city of Madrid and shows

the average price across neighborhoods in the city. These maps demonstrate that locations vary significantly in their capital prices. Finally, Figure 8 in Appendix B.3 shows the evolution of average time to sell over time, which was around 6 months during the boom of the real estate market and then increased to more than 10 months during the subsequent bust.

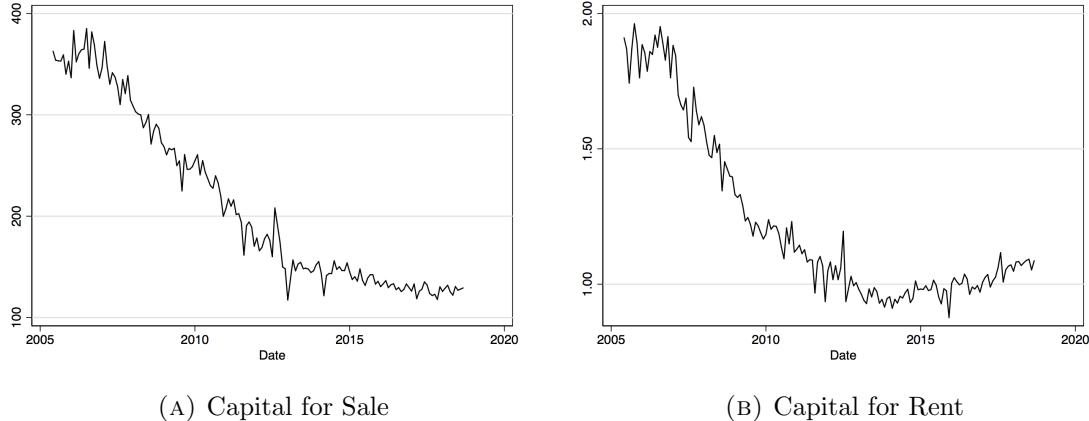
FIGURE 4: Capital Prices Across Locations



Note: Each map shows average prices by location expressed in constant 2017 dollars per square foot. The top panel shows average prices across provinces in Spain. The lower-left panel zooms in on the province of Madrid to show substantial heterogeneity across municipalities within this province. The lower-right map shows that, after zooming in on the municipality of Madrid, there is still significant geographical dispersion of prices across neighborhoods.

Discussion of the Data. The dataset has many advantages. First, it contains panel data for a large amount of nonresidential real estate, with wide geographical and temporal coverage. Second, it contains information on the duration of a listing online. Third, it provides information about the search behavior of potential buyers (monthly number of clicks and emails received). However, the dataset does not contain information on transacted prices. We believe that this should not be a concern for various reasons. First, Figure 5 in the Appendix compares indices of *listed* prices from Idealista with indices of *transacted* prices from the National Registry of Property in Spain. We

FIGURE 5: Evolution of Prices of Capital Units



Note: The left panel shows the evolution of mean prices at the daily frequency from 2006 to 2017. The right panel shows an equivalent index for rental units. Prices are denominated in constant 2017 dollars per square foot. To compute these price indices, we averaged the prices of all active listings in a given day.

show that the indices have similar patterns.¹² Second, below we show that listed prices are strongly associated with duration on the platform and the attention the listing receives (measured by clicks and emails received). Third, previous papers with access to both listed and transacted prices have shown that the modal property sells at its listed price and that the average property sells within 1.6% of its listed price (see, e.g., [Guren, 2018](#)).

Another concern is that duration contains measurement error due to sellers' failure to delete the listing after a sale. This should not be a concern. First, Idealista is a paid service, so it is costly for the seller to keep a listing dormant after the property has been sold. Second, a large fraction of listings are associated with professional sellers (i.e., real estate agents). Finally, to alleviate any remaining concerns, we exploit the fact that the platform asks sellers why they decided to close the listing. Figure 7 in Appendix B.3 compares the histogram of duration for two groups of listings: those that closed the listing because the property was rented out or sold and those that do not provide any explanation. Those histograms are virtually identical.

3.2 Key Data Moments

We now use our data to test the main predictions of the model. The theory predicts that if the quality of properties is known to all market participants, then their price should be negatively correlated with duration on the market. In addition, the theory predicts that if instead the quality of properties is private information of the seller, then their price should be positively correlated with duration on the market. In order to conduct these tests, we must first isolate the component of

¹²Our index leads the index of transacted prices. This is expected since our index consists of listed prices and it will take properties some months to exit the database, be registered as sales and recorded in national statistics.

a property's price that reflects the characteristics that are public information from the component that reflects the characteristics not observed by the econometrician, and potentially by buyers. Thus, we proceed by first measuring the component of a listed price that can be predicted based on the property's characteristics included in the listing. That is, we estimate hedonic regression to obtain the predicted price based on observable characteristics and the residual price. Then, we analyze how both the predicted and the residual price comove with duration on the market.

3.2.1 Obtaining a Measure of Residual Prices

To quantify the role of observable characteristics of a listing in explaining its price, we estimate the following hedonic pricing model for the (log) price per square foot:

$$\log(q_{ilt}) = \nu_{lt} + \gamma X_i + \varepsilon_{ilt} \quad (12)$$

where q_{ijt} is the price (in 2017 dollars) of a capital unit i in location l , listed in month t , ν_{lt} are location and time fixed effects, X_i is a set of observable characteristics included in the listing, and ε_{ilt} is a random error term.¹³ Table 2 presents the results of this exercise, showing the R^2 of the regression and the standard deviation of residual prices. To understand how much of the variation in prices is predicted by different groups of characteristics, we estimate multiple regressions including such groups one at a time.

Table 2 shows that in the raw data, the standard deviation of prices is 68% and 77% in the market for rent and sale, respectively. Of this variation, time fixed effects can account for 5% and 14% of the variation of prices in each type of transaction. This is perhaps surprising, given the collapse of the real estate market in Spain. However, this result is explained by the fact that in the real estate market, location is a key factor for determining the value of a property. Once we include location and time fixed effects, the R^2 of the regression increases to 53% and 59%, respectively. If we include interactions of the time-location fixed effects with the type of property (office, retail space, or warehouse), area, and age of the property, the R^2 further increases to 75% and 78%. Finally, in the last row of the table, we include the additional characteristics described in Table 1.¹⁴

¹³Location fixed effects are defined, for each unit, at the finest geographical level possible in the platform: the neighborhood level in the case of big cities like Madrid or Barcelona, and the city level in smaller cities. Results are similar if we focus only on cities that have available neighborhood information.

¹⁴In this analysis, we have focused on the average listed price during the lifetime of the listing. There is another source of price dispersion: price changes during the life of the listing. Table 1 in Appendix B.3 shows that between 5% and 7% of listings change price in a given month. Despite these price changes over time, most of the variation in prices across listings is accounted for by the average price of the listing. Table 2 in Appendix B.3 continues the analysis of Table 2 by further including a listing fixed effect and estimating the regression using the entire panel dataset. Results show that less than 6% of the variation in prices can be accounted for by properties that change their price during the lifetime of the listing.

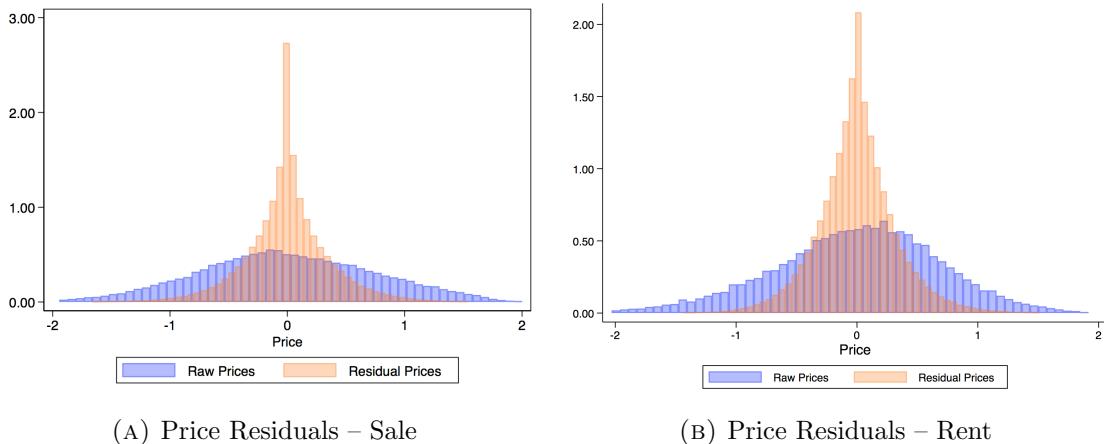
A major conclusion from Table 2 is that although the empirical model has a high predictive power for listed prices, between 20% and 25% of the variation in prices is not explained by characteristics included in the listings. Moreover, the standard deviation of the residuals in the benchmark specification, in which we include all available controls, is around 45% of the variation observed in the raw data. Figure 6 shows the distribution of price residuals, which illustrates the relevance of the dispersion in prices not accounted for by the characteristics in the listings. We refer to the dispersion of the residuals from the regression of log prices on all fixed effects and controls of the characteristics in the listings as *residual dispersion*.

TABLE 2: Price Variation Accounted for by Listed Characteristics

	Std Rent	Std Sale	R^2 Rent	R^2 Sale
Raw data	0.68	0.77	0.00	0.00
Time	0.67	0.71	0.05	0.14
Time x Location	0.47	0.49	0.53	0.59
... x Type	0.45	0.48	0.57	0.61
... x Area	0.35	0.37	0.74	0.76
... x Age	0.34	0.36	0.75	0.78
Benchmark	0.34	0.36	0.75	0.78

Note: This table shows the R^2 and standard deviation of residuals of different variations of equation (12). Time and location are fixed effects. Type (office, retail space, or warehouse), area, and age are sets of fixed effects for each of these characteristics. The row labeled Raw Data presents statistics for the demeaned raw log prices. The following rows include the mentioned fixed effects in the regression. The last row includes additional controls for the variables listed in Table 1.

FIGURE 6: Distribution of Price Residuals



Note: This figure shows the differences in log prices per square feet with respect to its mean for the raw data and price residuals after including the fixed effects in Table (2). The left panel shows the distributions for sales and the right panel for rentals.

3.2.2 Relationship between Prices and Duration

The last part of the analysis consists of documenting that residual prices and predicted prices have relationships with duration of opposite signs. By *predicted prices*, we refer to the component of a property's price that is linked to its observable characteristics. Since such characteristics are observed in the listing, the theory predicts that on average, properties with better characteristics (which are reflected by a higher predicted price) should have a shorter duration on the market. By *residual prices*, we refer to the component of a property's price that is *cannot* be explained by its observable characteristics. These residual prices can reflect the quality that is private information of the seller, but can also reflect characteristics that are observed by market participants, but not by the econometrician (e.g., listings that include pictures of the property, which convey information not included in the hedonic regression). A priori, the theory does not offer a prediction about the relation between residual prices and duration. If most of the residual price is the result of characteristics not observed by the econometrician, then this relationship should be negative. If, instead, a large component of residual prices is the result of sellers' choices to signal their private information, we should expect a positive relationship.

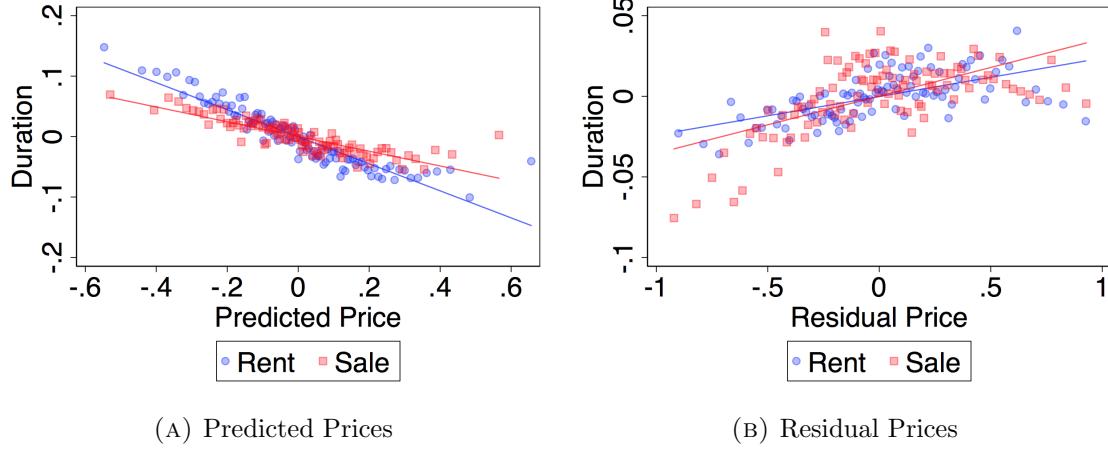
Figure 7 plots these relationships. Panel (a) shows that units with higher predicted prices tend to have a shorter duration on the market, which is consistent with model predictions under Full Information. Panel (b) shows that units with higher price residuals tend to have a higher duration, on the market. Table 3 presents the same results in a regression framework. In column (1), we regress (log) duration on (log) prices and obtain a negative and statistically significant relation. If the price of a property increases by 1%, expected duration decreases by 0.06%. In the second column, we split the (log) price into two components—predicted and residual prices—and run the same regression. While we obtain a negative and statistically significant relationship between duration and predicted prices, we obtain a positive and statistically significant relationship between duration and residual prices. In the last two columns, we split the analysis by type of operation (rent or sale), and results are similar across regressions.¹⁵

Table 5 and Figure 10 in Appendix B.3 reproduce the same analysis by replacing duration with the average monthly clicks received by a listing (as a proxy for search intensity). Results are consistent with those found for duration. Properties with high predicted prices receive more clicks on average, which is consistent with a shorter duration. On the other hand, properties with high

¹⁵The reason behind the inclusion of time-location fixed effects in the regression is to allow for the process of duration on the market to differ over time and location (e.g., the match efficiency could be market specific). However, the theory predicts that if a better observable location contributes positively to the quality of the property, it should also have a positive effect on the trading probability. Therefore, the inclusion of fixed effects is absorbing part of this effect as well. Tables 3 and 4 in the Appendix show the results excluding the location-time fixed effects. Results are robust to the exclusion of fixed effects.

residual prices receive fewer clicks on average, which is consistent with a longer duration. This last set of results is important, because it shows that listed prices do play an important role in attracting or repelling potential buyers by affecting their search behavior.

FIGURE 7: Relationship between Prices and Duration



Note: Panel (a) shows the relationship between predicted prices and duration. Panel (b) shows the relationship between residual prices and duration. Price residuals and predicted prices are obtained after running a regression of log prices on a set of fixed effects and observable characteristics. Figures show a binned scatter plot of each relationship, after controlling for location-time-type (offices, retail space, and warehouses) fixed effects.

TABLE 3: Prices and Duration

	(1)	(2)	(3)	(4)
	log(Dur)	log(Dur)	log(Dur)	log(Dur)
log(Price)	-0.068*** (0.001)			
log(Predicted Price)		-0.137*** (0.006)	-0.230*** (0.009)	-0.118*** (0.009)
log(Residual Price)		0.033*** (0.003)	0.029*** (0.004)	0.038*** (0.004)
Constant	2.014*** (0.002)	2.143*** (0.011)	1.801*** (0.001)	2.528*** (0.042)
Observations	1090875	1090857	643713	447117
R ²	0.010	0.202	0.199	0.262
Fixed Effects	Yes	Yes	Yes	Yes
Sub Sample	Full	Full	Rent	Sale

Note: This table presents the results of a regression of log duration on the two components of prices, residual and predicted prices. Because the left-hand-side variable is the log duration of a listing, we choose the right-hand-side variable to be the mean price over the lifetime of the listing. The first column shows a regression of duration on prices. Column 2 regresses duration on predicted, residual prices, and location×time×type fixed effects. Columns 3 and 4 split the sample by operation (sale versus rent). Standard errors are clustered at the location-time level. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

The different relations that residual and predicted prices have with duration suggest an important role of information. When higher prices stem from listed characteristics, such as the location of the unit—which can be perfectly observed by buyers—they tend to be associated with shorter

time to sell. When high prices cannot be easily linked to observable characteristics, then they are associated with longer time to sell.

Under the null hypothesis of full information, according to our model, residual prices reflect characteristics of properties not observed by the econometrician. Then, we should expect a negative relation with duration, as is the case with predicted prices. The fact that we estimate a positive relation provides evidence that the extent of asymmetric information cannot be zero. This conclusion is more formally supported in the estimation exercise of the model, which allows us to provide a quantitative magnitude of the deviation from full information.

3.2.3 Additional supporting evidence

3.3 Alternative explanations

In this subsection, we discuss alternative theories that could generate a positive relation between residual prices and duration. In each case, we present evidence showing that such alternative explanations are implausible, either because they have trouble reasonably matching the magnitude of the relation between residual prices and duration, and/or because they fail to simultaneously rationalize the relation between predicted prices and duration.

3.3.1 Search theories of price dispersion

The positive relation between residual price and duration suggests a trade-off between price and time-to-sell, giving rise to a natural explanation from the search literature. If properties and agents are homogeneous, price dispersion obtains from sellers' indifference when choosing the price of their listed units: Higher prices are associated with less search from buyers and more time to sell, and lower prices are associated with more search and shorter time to sell. The key is that the trade-off between prices and time to sell is such that they provide an equivalent expected revenue to the seller. This explanation is akin to that of labor- and product-market models such as [Burdett and Judd \(1983\)](#) and [Burdett and Mortensen \(1998\)](#).

To analyze the possibility that such theories explain the positive relation between residual prices and duration, we exploit the data to compute the expected net present values of listed properties at their residual prices and trading probabilities implied by the relation in Panel (b) of Figure 7. That is, we compute

$$\frac{pq}{(1 - \beta(1 - p))},$$

where q is the residual price, p is the selling probability, and β is the discount factor. The formula is just a geometric sum corresponding to the expected net present value of listing a price q and

with a selling probability p . Note that we abstract from price changes in this calculation and use the mean price instead, since the frequency of price changes is small.¹⁶

FIGURE 8: Net Present Value of Price-Duration Trade-off

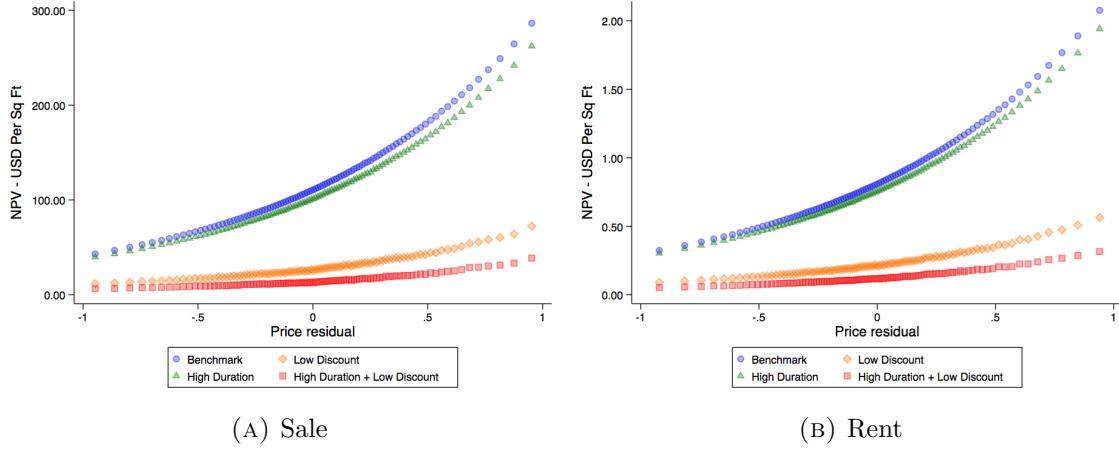


Figure 8 shows the results for both types of transactions. The blue circle lines correspond to our benchmark calculation, in which we use the observed duration and mean listed prices, and a discount factor of $\beta = 0.99$ to compute the net present value. The data show that models of frictional price dispersion cannot explain the positive relationship between residual prices and duration. In the data, the relationship is such that sellers cannot be optimally choosing to randomize: The expected net present value is monotonically increasing in the listed price. Any seller facing such price-duration trade-off will maximize expected revenue by choosing the highest price we see in the data.

3.3.2 Heterogeneous Sellers

Another potential explanation for the positive relation is that residual prices do not reflect heterogeneity in the properties' unobserved characteristics, but rather heterogeneity in sellers' preferences. Before we computed the net present value of a listing for a risk-neutral and relatively patient seller (which makes the preference for a higher price stronger). Here, we explore how our previous conclusion is affected by different preferences.

The orange diamond series in Figure 8 corresponds to an alternative calculation in which we make sellers extremely impatient ($\beta = 0$). The rationale for this calculation is that we may be ignoring heterogeneity in sellers' discount factors which leads us to conclude that some properties have higher returns when in reality low-price sellers are setting lower prices in order to sell faster, given their low discount factor. By setting the discount factor to 0, the calculation disproportionately

¹⁶The trading probability in a given month is computed from the duration of each property by $q = 1 - e^{-\lambda}$, where $\lambda = 1/duration$ is the hazard rate.

affects properties that have lower trading probabilities and high prices, flattening the net present value profile. However, even in this extreme scenario, the relation between prices and duration in the data is such that we still find that the net present value is monotonically increasing in the listed price. If, under the preferences of the most impatient seller, a higher residual price with lower selling probability is preferred, then the higher residual price would also be preferred under any other possible discount factor.

The green series with triangles shows a case in which the realization of duration is worse ex ante than the one that is realized ex-post. The rationale for this exercise is as follows. If sellers are heterogeneous with respect to their risk aversion, then some sellers may post lower prices as a measure to insure themselves. To evaluate the quantitative effect of this argument, we compute the NPV under extreme risk aversion: Sellers form expectations of trading probabilities under a “worst-case” scenario. We create quantiles of the price residual, and within each quantile we compute the standard deviation of duration across listings. Then, to compute the trading probability of each property, we use the realized duration plus 2 standard deviations of the duration within the quantile to which each property belongs. If the distribution of durations is more dispersed (riskier) for higher prices, then this exercise will shift the net present value of more expensive properties, flattening the NPV profile. The green series in Figure 8 show that this is indeed the case, but the quantitative magnitude is small: The NPV profile is still upward sloping. Finally, the red square series combines both sources of seller heterogeneity: It computes the NPV with both a zero discount factor and an adverse duration realization. Even in this case, the NPV profile is upward sloping.

The conclusion of this analysis is that seller heterogeneity cannot rationalize the positive relationship between residual prices and duration. If it could, the NPV analysis should show that sellers with different preferences should have different NPV-maximizing prices. We find that for a very broad set of preferences (discount factors from 0 to 0.99, and attitude toward risk from risk neutral to extreme forms of risk aversion), all sellers would maximize their expected net present value by choosing the highest price observed in the data.

The last piece of evidence against the role of sellers’ preferences is the negative relation between the predicted prices and duration we estimate. According to Proposition 2 in our model, sellers with homogeneous properties but different discount factors will optimally choose different prices: The optimal price is increasing in the seller’s discount factor, since a higher price is associated with a lower trading probability, which is relatively less costly for more patient sellers. Importantly, this prediction applies in the case with full information. If predicted prices reflect the component of a property’s quality that is observable to market participants, then a model with heterogeneous sellers’ discount factor should predict a positive relation between predicted prices and duration.

However, we instead estimate a negative relationship. Our claim is not that heterogeneity in sellers' preferences is not important, but that such heterogeneity cannot be the main driver of the data.

3.3.3 Heterogeneous Holding Costs

There is one additional source of sellers' heterogeneity that is not considered by our previous analysis of preference heterogeneity: the presence of heterogeneous holding costs. These are understood as costs that sellers must pay each period until the property is sold, such as maintenance costs, taxes, debt service costs, etc. If sellers face different costs, then some sellers might be forced to list properties at low prices in order to sell their property faster, as would occur in a fire sale. This would generate the positive relation between residual prices and duration. We explore this possibility by computing the size of this cost that would render sellers indifferent between listing at the highest price without a cost or at their chosen price with a cost. That is, we compute the necessary holding cost of each listing relative to the holding cost of a seller who listed his property at the highest price. We then present the size of this cost as a share of the listed price. The question we seek to answer with this exercise is: How large must the cost be in order to rationalize the choice of a lower residual price?

To compute the (unobserved) cost, c , we solve the following equation:

$$\frac{p_h q_h - c_j(1 - p_h)}{1 - \beta(1 - p_h)} = \frac{p_j q_j - c_j(1 - p_j)}{1 - \beta(1 - p_j)},$$

where q_h and p_h are the price and selling probability of the property with the highest residual price, respectively, and q_j and p_j are the price and probability we observe for property j . In order to rationalize the preference for a lower price and a higher trading probability, the cost must be higher when the difference between prices is larger and when the difference in durations is smaller. We estimate how large these costs must be in order to rationalize the choice of sellers for the case of sales. We present the cost normalized by the price of the most expensive property (in relative terms). Figure 9 presents the results. We can see that in order for differential holding costs to explain the differences in returns in the data, they must be extremely large. To illustrate, the cost of holding 1 square foot of a property for 1 additional month would have to be larger than the price at which the owner can sell that unit. We conclude that it is unlikely that the bulk of the positive relation between residual prices and duration is explained by the presence of heterogeneous holding costs.

FIGURE 9: Required Holding Costs



4 The Relevance of Information Frictions in the Capital Market

In this section, we calibrate the model to quantify the extent of asymmetric information in the market for physical capital. We then explore the effects of information frictions on allocations, prices, and liquidity by comparing equilibria with asymmetric and full information.

4.1 Quantifying Frictions: Model and Data

Calibration We calibrate the model under the assumption that the data are generated from a separating equilibrium in steady state. We choose to calibrate the model to a monthly frequency. There are two groups of parameters. The first group is set outside of our calibration exercise: β , η , χ , φ_K , and φ_E . The subjective discount factor β is chosen to match an annualized interest rate of 4%. The elasticity of the matching function η is set to 0.86 following [Ottonello \(2017\)](#), who estimates it using data from Idealista. Without loss of generality, we normalize the search cost χ to one.¹⁷ We set φ_E to 0.008 to match an annual firm exit rate of 9% (EUROSTAT). The assumption here is that when firms exit, they list their commercial real estate for sale. Finally, we set the number of types z and a to 100 each. These parameters are summarized in Table 4.

We jointly calibrate a second group of parameters— δ , σ_a^2 , σ_z^2 , and \bar{m} —by matching important identifying moments in the data. For this, we choose a set of moments as targets of the calibration of the model: the standard deviation of log residual prices, the covariance between residual prices and duration, the unconditional average duration, and the coefficients of a regression of log prices on the characteristics of properties (see equation (12)). For a given set of parameters, we compute

¹⁷We can show the following result. Take any combination (\bar{m}_1, χ_1) . Then, for any given value χ_2 , the equilibrium allocation, prices, and transaction probabilities with $\bar{m}_2 = \left(\frac{\bar{m}_1}{\chi_1^{1-\eta}}\right) \chi_2^{1-\eta}$ are the same.

TABLE 4: Externally Set Parameters

Parameter	Description	Value
β	Monthly discount factor	0.997
η	Elasticity matching function	0.860
χ	Search Cost (normalized)	1.000
$\varphi_K = \varphi_E$	Transition prob.	0.008
N_z	Number of z types	100
N_a	Number of a types	100

the equilibrium choices of prices and transaction probabilities for each type of capital. Then, we simulate the evolution of multiple units of capital, generate a sample of listed units as in Idealista, and perform the same analysis as the one performed with the microdata from Idealista. Finally, we compute those identifying moments with the simulated data and use a minimum-distance estimator to choose parameter values that match the moments in the data. Since the model is meant to capture the equilibrium within a specific market, we choose to calibrate it to a neighborhood in the city center of Madrid. Appendix C provides more details of this procedure.

Although there is no one-to-one mapping from parameters to moments, we provide intuition of the identification of the model parameters. The average transaction probability is pinned down by the matching efficiency \bar{m} . The vector δ transforms the characteristics of a property into the dividend it produces. Since dividends are reflected in the price of a property, those parameters are identified by the parameters of the hedonic regression of prices on the characteristics of units of capital. Finally, the key parameters σ_a^2 and σ_z^2 are jointly identified by the standard deviation of residual prices (those obtained after estimating the hedonic regression) and the covariance of residual prices and duration. The intuition is that the dispersion of prices is increasing in both σ_a^2 and σ_z^2 . However, as we showed before, the covariance responds differently to changes in σ_a^2 than to changes in σ_z^2 . The higher σ_a^2 is, the larger the component of residual prices driven by quality that is private information of the seller. According to the predictions of the model, there should be a more positive covariance between residual prices and duration. On the other hand, the higher σ_z^2 is, the larger the component of residual prices driven by observable quality. In this case, there should be a more negative covariance between residual prices and duration. These two opposing forces allow us disentangle the relative magnitudes of σ_a^2 and σ_z^2 .

Table 5 reports the calibrated parameters. The value of the matching efficiency implies that on average, properties stay on the market for 10 months. The presence of large trading frictions allows sellers of high-quality capital to signal their type by spending more time on the market. The values of σ_z^2 and σ_a^2 imply large quality heterogeneity across units of capital. Without including variation in observable characteristics, the standard deviation of (log) quality is 0.67. Of this overall

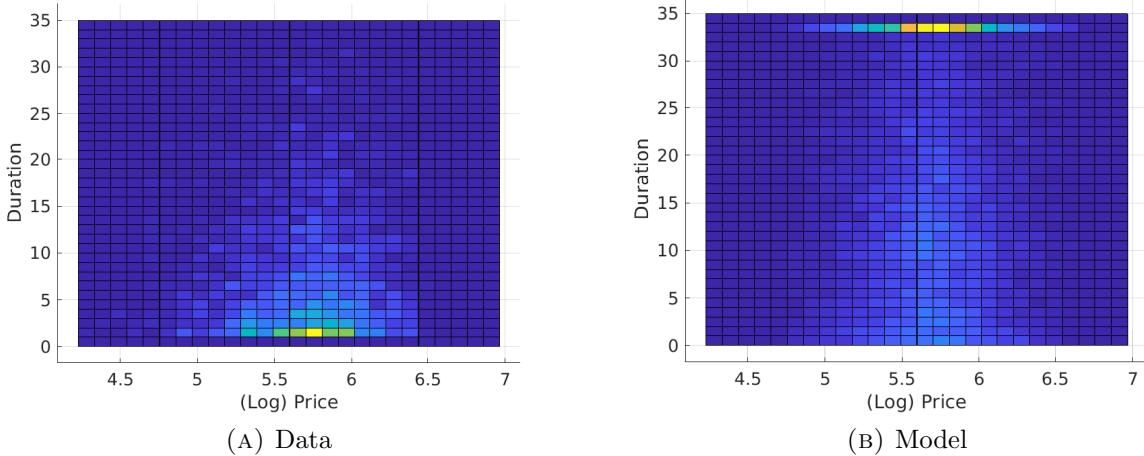
heterogeneity, 21% is attributed to heterogeneous quality that is private information of the seller.

TABLE 5: Internally Calibrated Parameters

Param.	Description	Value	Target	Data	Model
\bar{m}	Match efficiency	1.050	Avg. log duration	1.802	2.523
σ_z^2	Var. z	0.280	Var. log prices	0.124	0.122
σ_a^2	Var. a	0.030	Var. log duration	0.643	1.065
σ_{za}	Cov. z and a	0.000	Cov. log prices and duration	0.017	0.012

Goodness of Fit Figure 10 presents the joint distribution of prices and duration, in both the data and the model. The overall fit is good. Although the model generates excess mass at high durations, the distribution of prices is well approximated (additional results on the goodness of fit are presented in Appendix C). Additionally, in Table 6 we reproduce the regression results reported in Table 3 for this specific market. There is a positive relationship between residual prices and duration. This is expected, since we targeted the covariance between these two variables. However, the regression results also report a negative relationship between predicted prices and duration, in both the data and the model. This is an additional test of the quantitative exercise, since it was not part of the set of targeted moments.

FIGURE 10: Joint Distribution of (log) Prices and Duration



Note: These figures plot the joint distribution of (log) prices and duration in the data (Panel (A)) and in the model (Panel (B)). Lighter colors denote a higher concentration of observations in given price-duration pair.

5 Counterfactual Analyses

In this section, we perform a series of counterfactual analyses to quantify the effects information frictions have on allocations, prices, and liquidity. We begin by deriving a set of equilibrium objects

TABLE 6: Regression coefficients

	Data $\log(Dur)$	Model $\log(Dur)$
Intercept	3.119 (0.579)	2.780 (0.314)
Predicted price	-0.231 (0.102)	-0.045 (0.055)
Residual price	0.141 (0.040)	0.097 (0.021)
N	3247	19332
R^2	0.005	0.001

Note: This table presents the results of a regression of log duration on the two components of prices, residual and predicted prices. Predicted and residual prices are obtained from a first regression of (log) prices on the observable characteristics of capital (denoted by the vector X). The first column estimates the regression with data from Idealista. The second column estimates the regression with simulated data generated from the model.

that will guide the discussion of the role of asymmetric information.

Capital Unemployment As previously shown, information frictions generate lower average trading probabilities. Therefore, one of the variables affected by asymmetric information is the stock of unemployed capital. To derive the implications for capital unemployment, we begin with the law of motion of the stock of unemployed capital of quality (ω, a) :

$$k'_K(\omega, a) = k_E(\omega, a)\varphi_K(1 - p(\omega, a)) + k_K(\omega, a)(1 - \varphi_E)(1 - p(\omega, a)).$$

The stock of unemployed capital in the following period contains: (1) units of capital that are employed in the current period, whose owner experiences a negative shock and becomes a capitalist, and are not sold at the beginning of the following period, and (2) units of capital that are unemployed in the current period, whose owner does not become an entrepreneur, and are not sold at the beginning of the following period. The steady state unemployment rate thus becomes:

$$u(\omega, a) = \frac{\varphi_K(1 - p(\omega, a))}{p(\omega, a) + (\varphi_K + \varphi_E)(1 - p(\omega, a))}.$$

Finally, we can compute the steady state aggregate stock of unemployed capital as

$$\mathbb{E}(u(\omega, a)\omega a) = \underbrace{\mathbb{E}(u(\omega, a))}_{\text{Avg. unemployment rate}} \mathbb{E}(\omega a) + \underbrace{\text{Cov}(u(\omega, a), \omega a)}_{\text{Composition of unemp. capital}}.$$

Asymmetric information affects the stock of unemployed capital via two effects. First, asymmetric information reduces average trading probabilities, and therefore increases the average unemploy-

ment rate across units of different qualities. Second, there is a composition effect. Because asymmetric information reduces more the trading probabilities of units of capital of higher quality, the pool of unemployed capital is composed of units with higher average quality than the unconditional average quality in the economy.

Illiquidity Discounts The reduction in trading probabilities caused by asymmetric information has implications for the valuation of assets. To see this, recall that the price of an asset of quality (ω, a) is given by

$$p(\omega, a) = v_{\mathcal{E}}(\omega, a) - \frac{\chi}{\mu(\theta(\omega, a))},$$

thus reflecting the valuation of the asset by the entrepreneur and the total search costs incurred to purchase the unit of capital. In turn, we can express the value of the unit of capital to the entrepreneur as

$$\begin{aligned} v_{\mathcal{E}}(\omega, a) &= \frac{\omega a}{1-\beta} \left(1 - \underbrace{\frac{\beta \varphi_{\mathcal{K}} (1-p(\omega, a))}{(1-\beta)(1-\varphi_{\mathcal{K}}-\varphi_{\mathcal{E}})(1-p(\omega, a))}}_{\text{Illiquidity discount}} \right) \\ &\quad - \underbrace{\frac{\beta \varphi_{\mathcal{K}} \chi \theta(\omega, a)}{(1-\beta)(1-\beta(1-\varphi_{\mathcal{K}}-\varphi_{\mathcal{E}})(1-p(\omega, a)))}}_{\text{Search costs}}. \end{aligned}$$

As in [Duffie et al. \(2005\)](#), the fact that current entrepreneurs might become sellers in the future means that the value of a unit of capital for an entrepreneur today is affected by trading frictions. First, the value of capital includes the net present value of dividends, reduced by an illiquidity discount. Entrepreneurs value less units of capital that take longer to sell. Second, since future sale prices are discounted by buyers' total search costs, a lower buyers' trading probability increases such costs and reduces the sale price. Therefore, a higher market tightness $\theta(\omega, a)$ decreases the illiquidity discount, but increases buyers' search costs.

Counterfactual 1: Quantifying the Effects of Asymmetric Information Having estimated the parameters of the model, we quantify the effects of asymmetric information on equilibrium allocations, prices, and liquidity. To do so, we compare the model's predictions with a counterfactual scenario in which there is no private information. The latter corresponds to a version of the model in which the a component of quality is perfectly observed by all market participants.

Results are reported in Table 7. The first column reports equilibrium objects in the estimated model and the second column reports them for the counterfactual scenario with full information. Of the aggregate stock of efficiency units of capital of 1.3 (the stock of units of capital is normalized to

1), 0.24 are misallocated due to information frictions. The average unemployment rate of capital is 18% with asymmetric information, and close to zero with full information. The contribution of the composition effect to capital unemployment is small: The stock of unemployed capital (in efficiency units) is only 4% higher, because higher-quality capital is more likely to be unemployed. These results are further illustrated by the second panels in Figures 11 and 12, which show the equilibrium trading probabilities of properties with different (z, a) qualities. In the case with asymmetric information, the gradient of the trading probability with respect to the unobserved quality a is negative and steep. Instead, with full information, this gradient is positive and not as steep.

Table 7 further reports that information frictions decrease the unconditional average price of units of capital by 16.7% relative to the scenario with full information. The reduction in prices is almost entirely brought about by a decrease in the entrepreneurs' valuations of assets (the contribution of the reduction of buyers' search costs is negligible). In turn, entrepreneurs' valuations are severely affected by the illiquidity discount: Asymmetric information leads to an illiquidity discount of 16.4%. With full information, this discount is close to zero. These results are further illustrated in the first panels in Figures 11 and 12, which show the equilibrium prices of properties with different (z, a) qualities.

Finally, Table 7 presents the comparison of welfare across scenarios. In our model, welfare is given by total dividends produced by employed capital net of the search costs buyers pay in order to search and match with sellers. The latter are determined by the search cost χ and the mass of buyers searching in the market: $\chi \mathbb{E}(\theta(\omega, a) k_{\mathcal{K}}(\omega, a))$. In the estimated model, the welfare losses generated by information frictions are equivalent to 18.4% of welfare in the counterfactual with full information. In the estimated model, search costs are negligible. The fact that sellers face low trading probabilities means that buyers are meeting sellers with high probabilities, so they do not incur large search costs. With full information, search costs are slightly larger, but still represent only 0.02% of total welfare. To sum up, we find that the effects of information frictions in the market for physical capital are large. The magnitudes of the effects on allocations and prices are on the same order of magnitude of the estimated effects of trading frictions (e.g., Gavazza, 2016).

6 Conclusion

In this paper, we documented that information asymmetries play a key role in asset markets. This conclusion emerges from a new methodology to measure information frictions and from applying this methodology to microlevel data on the physical capital market. The methodology builds upon theories of asymmetric information in markets with trading frictions, which predict that information

TABLE 7: The Role of Information: Asymmetric vs Full Information

	Asym. Info.	Full Info.
$\mathbb{E}(u(\omega, a)\omega a)$	0.2419	0.0018
$\mathbb{E}(u(\omega, a))$	0.1786	0.0016
$\text{Cov}(u(\omega, a), \omega a)$	0.0090	-0.0003
$\mathbb{E}(q(\omega, a))$	318.7825	382.7327
$\mathbb{E}(v_{\mathcal{E}}(\omega, a))$	318.7829	382.9802
$\mathbb{E}(\chi/\mu(\omega, a))$	0.0004	0.2475
Avg. Illiquidity Disc.	0.1641	0.0016
Welfare	1.0624	1.3023
Output	1.0624	1.3026
Search Costs	0.0000	0.0003

Note: This table reports the results of counterfactual analysis. The first column reports statistics on capital unemployment, prices, and welfare in the calibrated model with asymmetric information. The second column reports similar statistics under the counterfactual scenario in which the a component of quality becomes the common knowledge of all market participants.

FIGURE 11: Prices and Trading Probabilities: Asymmetric Information

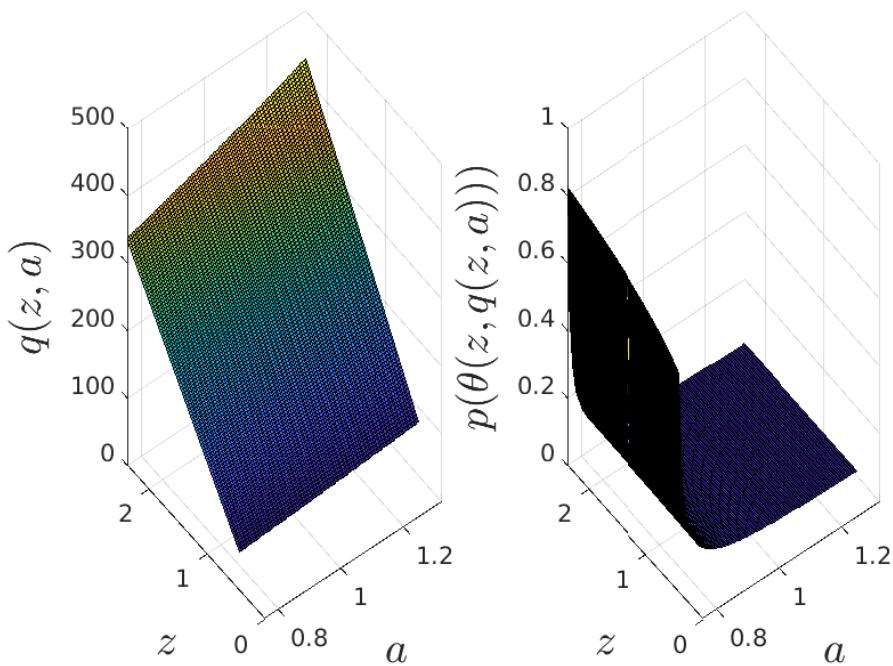
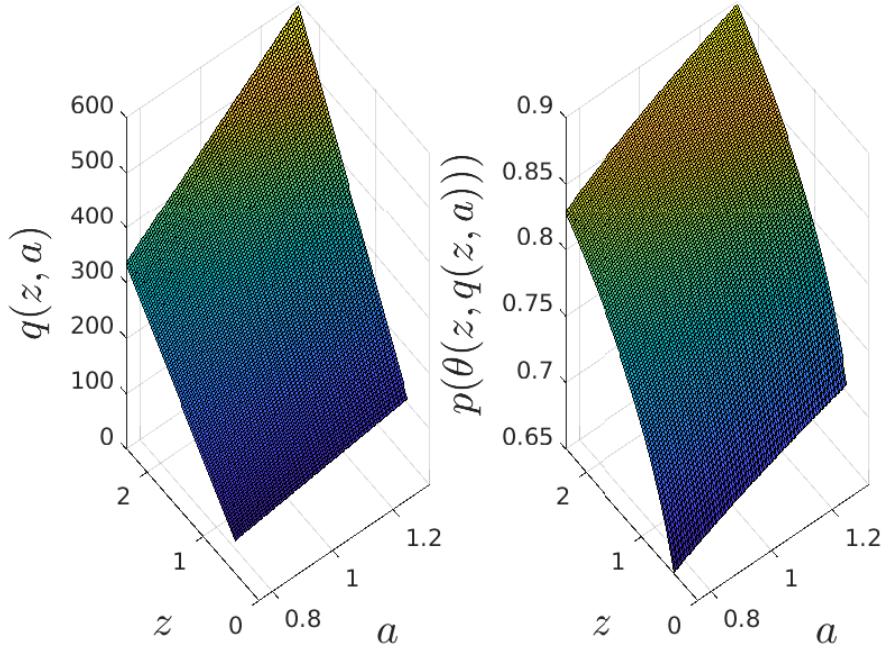


FIGURE 12: Prices and Trading Probabilities: Full Information



affects the relationship between prices and trading probabilities. The empirical analysis shows that the patterns between listed prices and duration are consistent with the presence of asymmetric information. Combining theory and data, we estimate that information frictions are large and have a significant effect on output, prices, and liquidity.

In future research, our methodology could be used to inform models of asymmetric information and business cycles. In this area, an important additional element to consider would be the presence of financial frictions that interact with information and trading frictions. For instance, if capital is used as collateral, asymmetric information could be a first order factor shaping firms' and households' access to financial markets. In addition, one could study how asymmetric information affects the production of new capital, and how policies can affect the choice of quality in aggregate investment. These extensions are planned for future research.

References

- AKERLOF, G. (1970): “The market for Lemons: Quality uncertainty and the market mechanism,” *Quarterly Journal of Economics*, 84, 488–500.
- BIGIO, S. (2015): “Endogenous liquidity and the business cycle,” *American Economic Review*, 105, 1883–1927.
- BURDETT, K. AND K. L. JUDD (1983): “Equilibrium price dispersion,” *Econometrica*, 955–969.
- BURDETT, K. AND D. T. MORTENSEN (1998): “Wage differentials, employer size, and unemployment,” *International Economic Review*, 257–273.
- CAO, M. AND S. SHI (2017): “Endogenously procyclical liquidity, capital reallocation, and q ,” Working paper.
- CAPLIN, A. AND J. LEAHY (2011): “Trading frictions and house price dynamics,” *Journal of Money, Credit and Banking*, 43, 283–303.
- CHIAPPORI, P.-A. AND B. SALANIE (2000): “Testing for asymmetric information in insurance markets,” *Journal of Political Economy*, 108, 56–78.
- CHO, I.-K. AND D. M. KREPS (1987): “Signaling games and stable equilibria,” *The Quarterly Journal of Economics*, 102, 179–221.
- DELACROIX, A. AND S. SHI (2013): “Pricing and signaling with frictions,” *Journal of Economic Theory*, 148, 1301–1332.
- DUFFIE, D., N. GÂRLEANU, AND L. H. PEDERSEN (2005): “Over-the-counter markets,” *Econometrica*, 73, 1815–1847.
- EINAV, L., A. FINKELSTEIN, AND P. SCHRIMPFF (2010): “Optimal mandates and the welfare cost of asymmetric information: Evidence from the uk annuity market,” *Econometrica*, 78, 1031–1092.
- EISFELDT, A. AND Y. SHI (2018): “Capital reallocation,” in *Annual Review of Financial Economics*.
- EISFELDT, A. L. (2004): “Endogenous liquidity in asset markets,” *The Journal of Finance*, 59, 1–30.
- EISFELDT, A. L. AND A. A. RAMPINI (2006): “Capital reallocation and liquidity,” *Journal of Monetary Economics*, 53, 369–399.

- GAVAZZA, A. (2011): “The role of trading frictions in real asset markets,” *The American Economic Review*, 101, 1106–1143.
- (2016): “An empirical equilibrium model of a decentralized asset market,” *Econometrica*, 84, 1755–1798.
- GAVAZZA, A. AND A. LANTERI (2018): “Credit Shocks and Equilibrium Dynamics in Consumer Durable Goods Markets,” Working paper.
- GUERRIERI, V. AND R. SHIMER (2014): “Dynamic Adverse Selection: A Theory of Illiquidity, Fire Sales, and Flight to Quality,” *American Economic Review*, 104, 1875–1908.
- GUERRIERI, V., R. SHIMER, AND R. WRIGHT (2010): “Adverse selection in competitive search equilibrium,” *Econometrica*, 78, 1823–1862.
- GUREN, A. M. (2018): “House price momentum and strategic complementarity,” *Journal of Political Economy*, 126, 1172–1218.
- HENDEL, I. AND A. LIZZERI (1999): “Adverse selection in durable goods markets,” *American Economic Review*, 89, 1097–1115.
- HOUSE, C. L. AND J. V. LEAHY (2004): “An sS model with adverse selection,” *Journal of Political Economy*, 112, 581–614.
- HSIEH, C.-T. AND P. J. KLENOW (2009): “Misallocation and manufacturing TFP in China and India,” *The Quarterly journal of economics*, 124, 1403–1448.
- IVASHINA, V. (2009): “Asymmetric information effects on loan spreads,” *Journal of Financial Economics*, 92, 300–319.
- KRAINER, J. (2001): “A theory of liquidity in residential real estate markets,” *Journal of Urban Economics*, 49, 32–53.
- KURLAT, P. (2013): “Lemons markets and the transmission of aggregate shocks,” *American Economic Review*, 103, 1463–89.
- KURLAT, P. AND J. STROEBEL (2015): “Testing for information asymmetries in real estate markets,” *The Review of Financial Studies*, 28, 2429–2461.
- KURMANN, A. AND N. PETROSKY-NADEAU (2007): “Search frictions in physical capital markets as a propagation mechanism,” Working Paper 07-12, Centre Interuniversitaire sur le Risque, les Politiques Économiques and l’Emploi.

- LAGOS, R., G. ROCHETEAU, AND R. WRIGHT (2017): “Liquidity: A new monetarist perspective,” *Journal of Economic Literature*, 55, 371–440.
- LANTERI, A. (2018): “The market for used capital: endogenous irreversibility and reallocation over the business cycle,” *The American Economic Review*, 108.
- LESTER, B., A. SHOURIDEH, V. VENKATESWARAN, AND A. ZETLIN-JONES (2018): “Screening and adverse selection in frictional markets,” *Journal of Political Economy*, forthcoming.
- MENZIO, G. AND S. SHI (2011): “Efficient Search on the Job and the Business Cycle,” *Journal of Political Economy*, 119, 468–510.
- MOEN, E. R. (1997): “Competitive Search Equilibrium,” *Journal of Political Economy*, 105, 385–411.
- MYERS, S. C. AND N. S. MAJLUF (1984): “Corporate financing and investment decisions when firms have information that investors do not have,” *Journal of Financial Economics*, 13, 187–221.
- OTTONELLO, P. (2017): “Capital Unemployment,” Working paper.
- PIAZZESI, M., M. SCHNEIDER, AND J. STROEBEL (2015): “Segmented housing search,” NBER working paper 20823.
- RAMEY, V. A. AND M. D. SHAPIRO (1998): “Capital churning,” Mimeo.
- (2001): “Displaced capital: A study of aerospace plant closings,” *Journal of Political Economy*, 109, 958–992.
- SHIMER, R. (2005): “The Cyclical Behavior of Equilibrium Unemployment and Vacancies,” *The American Economic Review*, 95, 25–49.
- SPENCE, M. (1973): “Job market signaling,” *The Quarterly Journal of Economics*, 87, 355–374.
- STIGLITZ, J. E. AND A. WEISS (1981): “Credit rationing in markets with imperfect information,” *The American economic review*, 71, 393–410.
- WHEATON, W. C. (1990): “Vacancy, search, and prices in a housing market matching model,” *Journal of Political Economy*, 98, 1270–1292.

A Theory Appendix

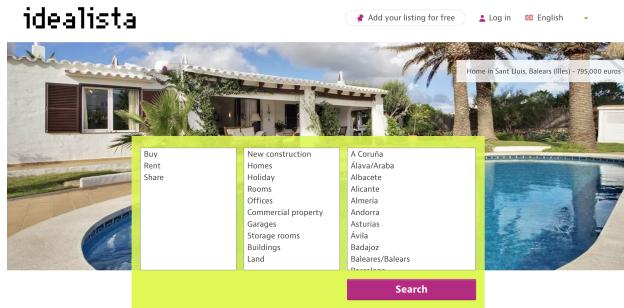
To be completed.

B Empirical Appendix

B.1 The online platform

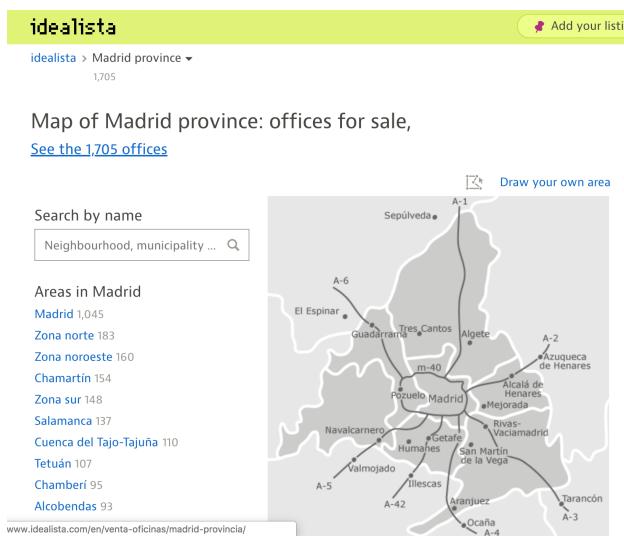
This subsection describes how the platform works. When entering the website, the buyer encounters the screen shown in Figure 1. The platform asks the client to choose a type of transaction (buy, rent, or find a shared space), the type of property (retail store, office, etc.), and the location.

FIGURE 1: Main Website



Once those options are selected, suppose the client wants to find a unit in Madrid (Figure 2). There, the website shows the number of properties available for sale by area in the city.

FIGURE 2: Options Madrid



After choosing a narrower location within the city (not shown here), the client finds a scrolling list of the available units that meet her requirements, as shown in Figure (3). There, the user can

include more filters depending on her requirements for layout and amenities.

FIGURE 3: Available listings in a narrow location in Madrid

The screenshot shows a search results page for 'Offices for sale in Prosperidad, Madrid' on the idealista website. The page header includes the date (10/4/2018), a navigation bar (idealista > Madrid province > Madrid > Chamartín > Prosperidad), and a user log-in option. The main content area displays 32 results, with the first three listed below:

- Office in CLARA DEL REY, Prosperidad, M...** **550,000 €** **2,709 eur/m²**
203 m² Magnificent office of 203m² built to reform, in building of 1980, in Prosperity area, with large areas and great
- Office in CLARA DEL REY, Prosperidad, M...** **1,250,000 €** **2,341 eur/m²**
510m² office to reform, in building of 1980, with two independent registration notes, in the Prosperidad area.
- Office in Clara del Rey, Prosperidad, Mad...** **700,000 €** **2,115 eur/m²**
331 m² Magnificent office of 331m² built to reform, in a building of 1980, in Prosperidad area, with large areas and great

Each listing includes a small thumbnail image of the interior, the price, the area per square meter, a brief description, and a contact phone number (918 004 132). The interface also features a 'Buy' tab, a 'Rent' tab, a 'Favourites only' filter, and various search filters on the left side.

When the user finds a unit that may be to her taste and clicks on it, a window pops up with the details shown in Figure 4 plus text details not shown here. The main information the listing contains is the unit description with pictures, price, change in price, area, date of construction, and other amenities and equipment.

FIGURE 4: A listing on the website

The screenshot shows a real estate listing for an office space. At the top, it says "Office for sale in CLARA DEL REY, Prosperidad, Madrid". The listing is from "idealista" and "ALFEREZ REAL ESTATE Madrid". It has a phone number "+34 918 004 132" and a reference "Ref.: AC-JFV-0134". There is a "Personal note" section where users can add a message. Below the header, there is a photograph of a hallway with wooden doors and yellow walls. The main listing details are as follows:

- Office for sale in CLARA DEL REY**
- Prosperidad, Madrid
- 550,000 €** ~~700,000 €~~ ↓ (21%)
- 203 m² | 2,709 eur/m²
- Basic features**
 - 203 m² built
 - Screen layout
 - 1 bathrooms within the office
 - Second hand/needs renovating
 - Built in 1980
- Building**
 - 1st floor exterior
 - 2 lifts
 - Mixed use
 - Doorman/guard
 - Security door
 - Fire extinguishers
 - Energy efficiency rating of the completed building: in progress
- Equipment**
 - Heating
 - Hot water
 - Air conditioning with cooling/heating function
 - Suspended ceiling

At the bottom, there is a URL "https://www.idealista.com/en/inmueble/82285673/" and a page number "1/2".

B.2 Representativeness of the dataset

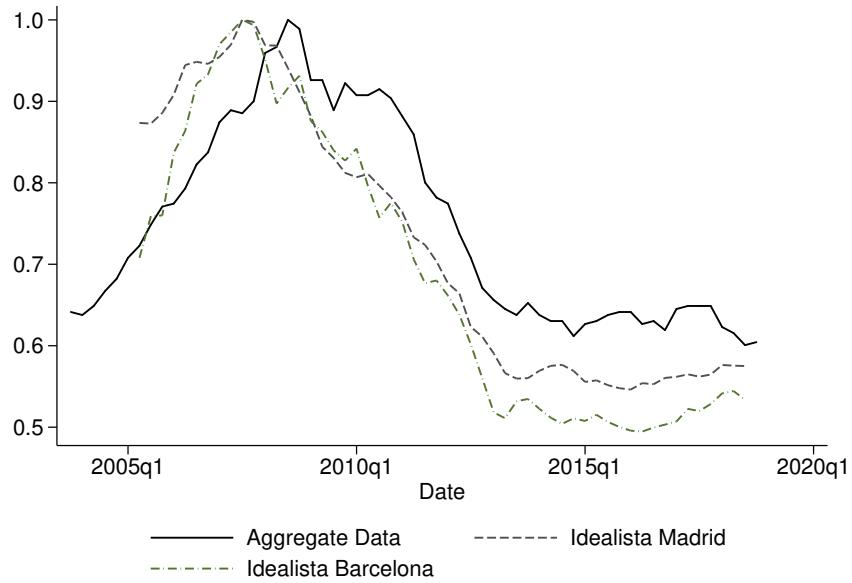
In this subsection, we analyze the representativeness of the dataset, showing that our data is consistent with aggregate patterns observed in Spain over this period. We provide two pieces of evidence about our data. First, we show that in our data, the price index exhibits the patterns of aggregate data. Second, we show that the patterns of sales follow those of aggregate sales of structures in Spain.

Figure 5 shows the index of listed prices for properties for sale in our sample and the index of transacted prices of retail space in Spain (the latter come from official transaction records). Both indexes are normalized to one at their respective peak. We highlight the fact that the fall in prices we observe is consistent, and very similar in size to that observed for retail space in Spain during the recent financial crisis. Moreover, our index leads the aggregate index, which is expected since our index consists of listed prices. This is expected, since our index consists of listed prices and it will take properties some months to exit the database, be registered as sales, and be recorded in national statistics.

Figure 6 shows the index of sales for properties for sale in our sample and the aggregate sales index of real estate in Spain. Both indexes are normalized to take the value of 1 in the first month of 2007. The index for our data is constructed by computing the share of units that exit the database

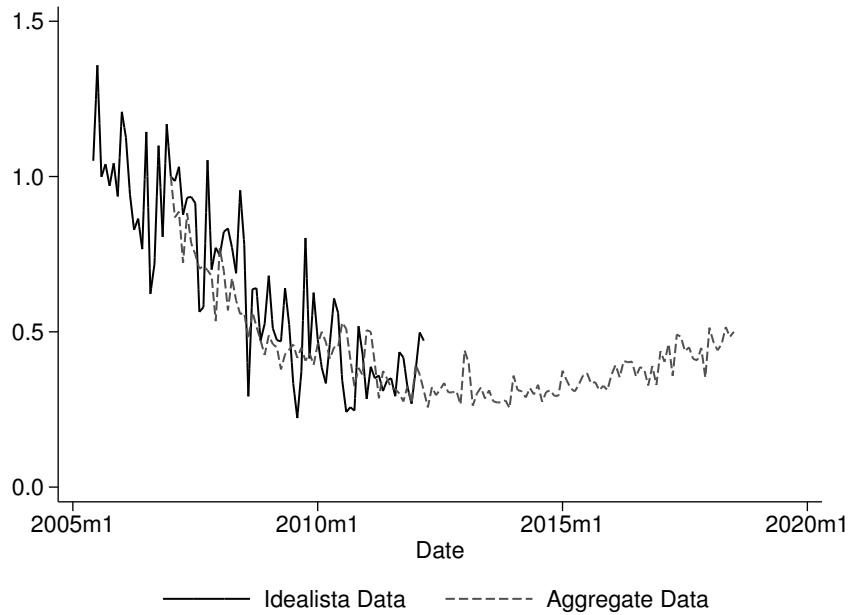
with respect to the number of active posts in that month. In the case of the aggregate number, we normalize the number of sale transactions recorded by the Statistical Agency. In doing this, we assume that the total stock of units during this period is fairly constant (we do not have information on the size of the stock). Although our index is noisier than the national estimates, the patterns of the two series are close to each other.

FIGURE 5: Price Index: Dataset versus Aggregate Data



Note: The solid line shows the price index for properties for sale in Barcelona and Madrid in our dataset. The dashed line shows the aggregate retail space price index gathered from the National Registry of Property (Registros de España). All indices are normalized to their respective peak.

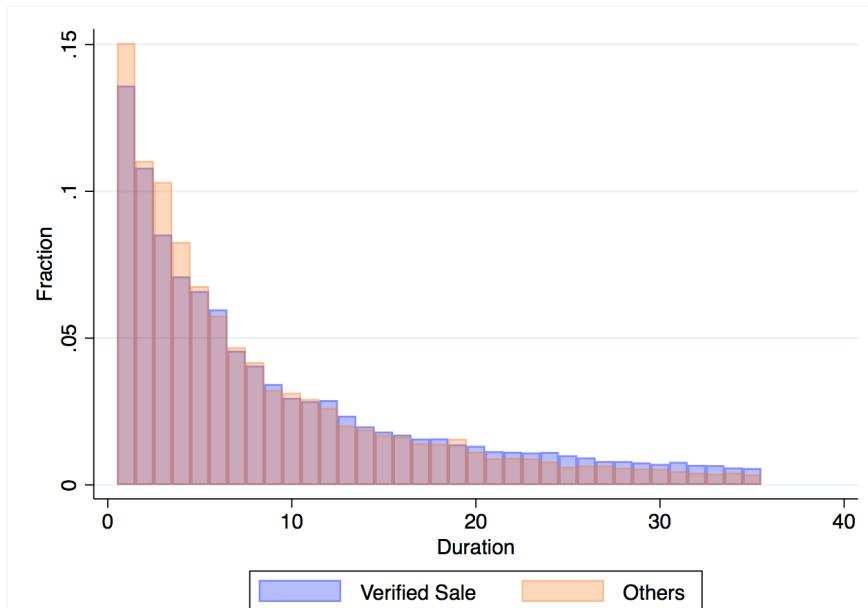
FIGURE 6: Sales Rate: Dataset versus Aggregate Data



Note: The solid line shows the sales rate for properties in our dataset. The dashed line shows the aggregate sales index of real estate gathered from the Statistical Agency of Spain (INE). Both indices take the value of one in January 2007.

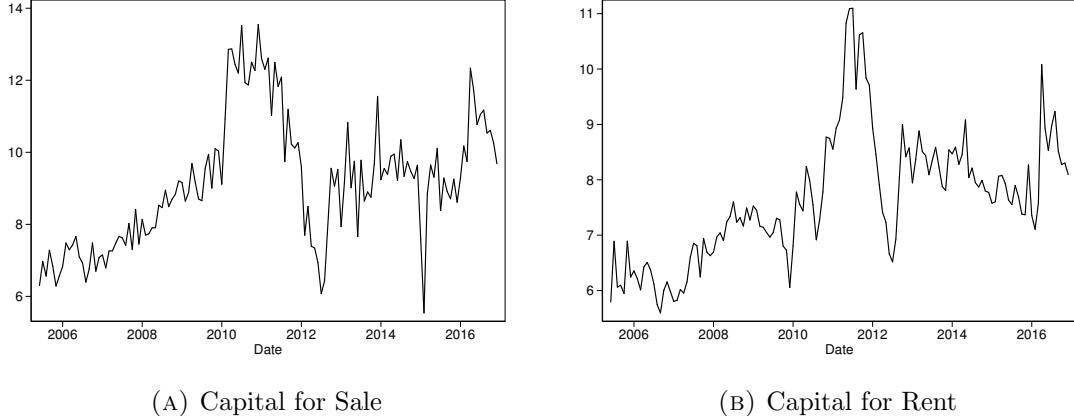
B.3 Additional Figures and Tables

FIGURE 7: Distribution of Duration: Confirmed Sales



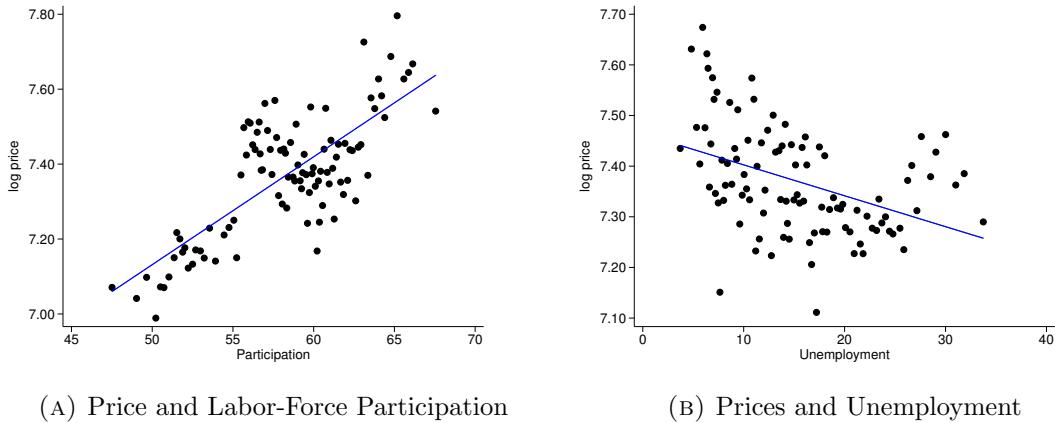
Note: This figure compares the histogram of duration for two subgroups of listings: those that, after removing the listing from the platform, explained that they did so because the property was rented out or sold, and those that did not provide an explanation.

FIGURE 8: Evolution of Average Duration



Note: The left panel shows the evolution of mean time to sell (in months) at monthly frequency from 2006 to 2017. The right panel shows an equivalent index for rental units. Time to sell is measured as the time difference between the entry and exit dates of each listing. Each observation contains the average time to sell for listings that entered the online platform in a given month.

FIGURE 9: Capital Prices and Regional Business Cycles



Note: This figure shows the relationship between log prices of posts for sale in a given province in a given quarter year with economic variables, in this case labor-force participation and unemployment rate. The figure presents a binned scatterplot, in which we choose 100 quantiles of the relationship between the economic variable of interest and log prices and compute the average for observations in that quantile.

TABLE 2: Price Variation Accounted for by Listed Characteristics in New Entrants

Statistic	Sale		Rent	
	IQR	R ²	IQR	R ²
Raw Data	0.666	0.000	0.802	0.000
Benchmark	0.284	0.776	0.198	0.845
Property Fixed Effect	0.119	0.946	0.116	0.937

Note: This table extends Table (2) by including a property fixed effect, which gathers inference from properties that change their prices while they are active in the dataset. We find that after including property fixed effects, non-parametrically absorbing all of the property's time-invariant price determinants, the IQR is roughly 11% and the R² is roughly 0.94.

TABLE 1: Frequency of Price Changes for Capital

	Rent Office	Sale Office	Rent Warehouse	Sale Warehouse
Frequency of Price Changes	0.07	0.07	0.05	0.07
Frequency of Price Increases	0.02	0.02	0.02	0.02
Frequency of Price Decreases	0.05	0.05	0.04	0.05
Absolute Size of Price Changes	0.15	0.12	0.16	0.15
Absolute Size of Price Increases	0.19	0.15	0.19	0.18
Absolute Size of Price Decreases	0.14	0.11	0.15	0.14

Note: This table presents price adjustment statistics by property type and operation. In order to compute the table, we first compute statistics about price changes within each property and then take averages across properties in a given time period. Finally, we compute the average over time. The first row shows the frequency of price changes, which is the average share of properties that exhibit a price change in a given month. The following two rows show the share of listings with price increases and decreases. The absolute size of price changes is computed as the absolute value of the log difference in prices over consecutive months (ignoring the zeros).

TABLE 3: Regression of Prices on Duration - Sale

	(1) log Dur	(2) log Dur	(3) log Dur
log(Price)	-0.012*** (0.002)		
log(Predicted Price)		-0.032*** (0.006)	-0.118*** (0.009)
log(Residual Price)		0.035*** (0.004)	0.038*** (0.004)
Observations	447141	447141	447117
R^2	0.000	0.001	0.262
Subsample	Full	Full	Full
Fixed Effects	No	No	Location-Time

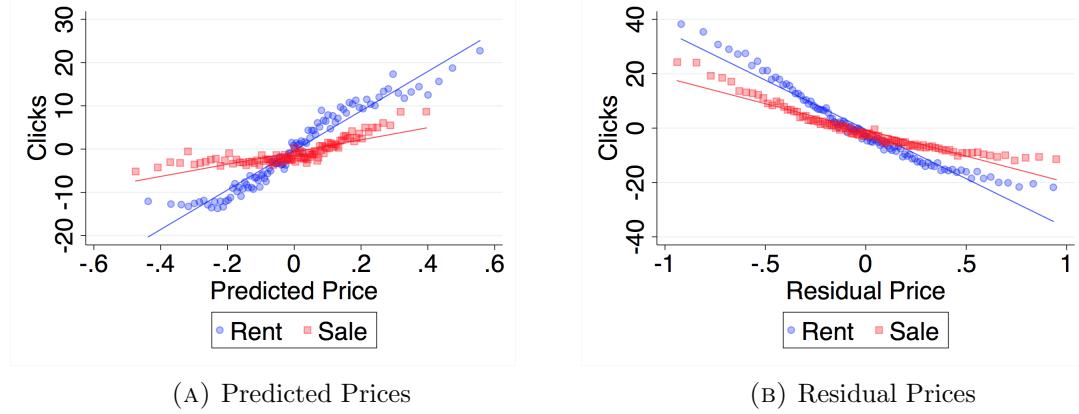
Note: This table presents the results of a regression of log duration on the two components of prices, residual and predicted prices. The sample includes listings for sale only. Because the left-hand-side variable is the log duration of a listing, we choose the right-hand-side variable to be the mean price over the lifetime of the listing. The first column shows a regression of duration on prices. Column 2 regresses duration on predicted and residual prices. Column 3 additionally includes location×time×type fixed effects. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

TABLE 4: Regression of Prices on Duration - Rent

	(1)	(2)	(3)
	log Dur	log Dur	log Dur
log(Price)	-0.116*** (0.002)		
log(Predicted Price)		-0.186*** (0.008)	-0.230*** (0.009)
log(Residual Price)		0.032*** (0.005)	0.029*** (0.004)
Observations	643734	643734	643713
R ²	0.009	0.015	0.199
Subsample	Full	Full	Full
Fixed Effects	No	No	Location-Time

Note: This table presents the results of a regression of log duration on the two components of prices, residual and predicted prices. The sample includes listings for rent only. Because the left-hand-side variable is the log duration of a listing, we choose the right-hand-side variable to be the mean price over the lifetime of the listing. The first column shows a regression of duration on prices. Column 2 regresses duration on predicted and residual prices. Column 3 additionally includes location×time×type fixed effects. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

FIGURE 10: Relationship between Prices and Clicks



Note: Panel (a) shows the relationship between predicted prices and average monthly clicks. Panel (b) shows the relationship between residual prices and average monthly clicks. Price residuals and predicted prices are obtained after running a regression of log prices on a set of fixed effects and observable characteristics. The figures show a binned scatter plot of each relationship, after controlling for location-time-type (offices, retail space, and warehouses) fixed effects.

C Quantitative Analysis

C.1 Model Solution and Simulation

To be completed.

TABLE 5: Prices and Clicks

	(1) Clicks	(2) Clicks	(3) Clicks	(4) Clicks
$\log(\text{Price})$	3.437*** (0.113)			
$\log(\text{Predicted Price})$		28.218*** (0.991)	41.899*** (1.683)	8.710*** (0.648)
$\log(\text{Residual Price})$		-31.318*** (0.828)	-40.048*** (1.102)	-19.907*** (0.521)
Constant	52.649*** (0.224)	6.464*** (1.845)	75.808*** (0.231)	1.649 (3.078)
Observations	1070976	1070810	632581	438064
R^2	0.028	0.369	0.399	0.341
Subsample	Full	Full	Rent	Sale

Note: This table presents the results of a regression of log average monthly clicks on the two components of prices, residual and predicted prices. Because the left-hand-side variable is the average monthly clicks a listing received, we choose the right-hand-side variable to be the mean price over the lifetime of the listing. The first column shows a regression of clicks on prices. Column 2 regresses clicks on predicted and residual prices, and location \times time \times type fixed effects. Columns 3 and 4 split the sample by operation (sale versus rent). Standard errors are clustered at the location-time level. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

C.2 Goodness of Fit

FIGURE 11: Regression Coefficients: Data vs Model

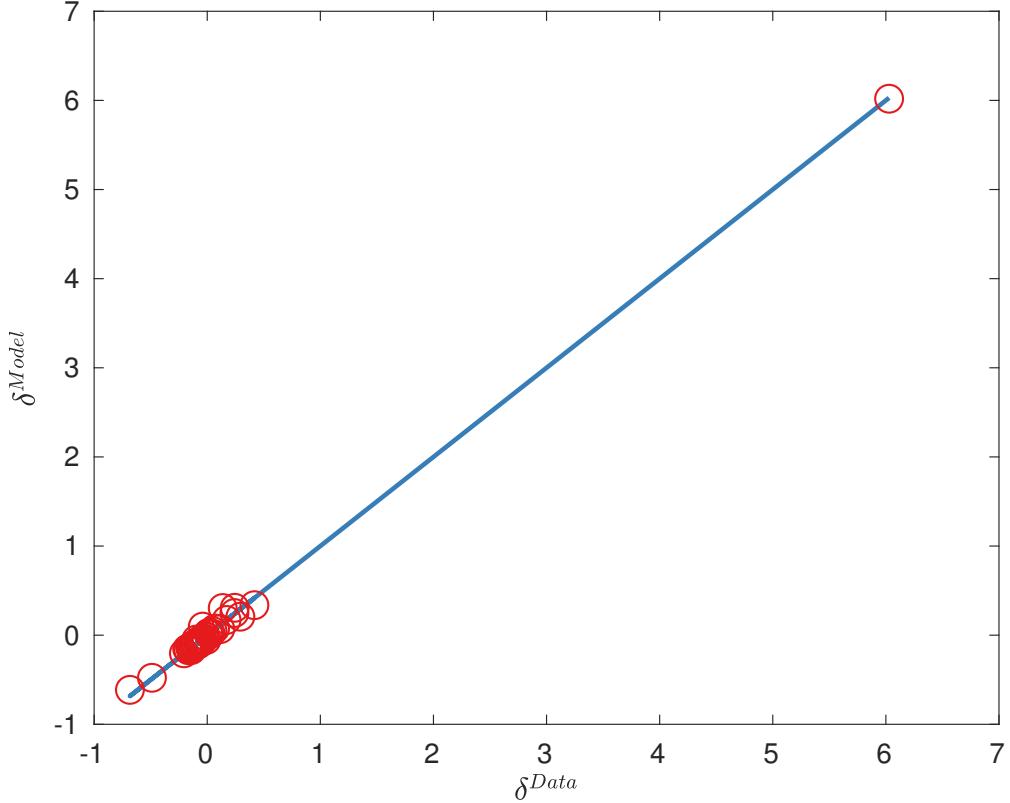


FIGURE 12: Histogram of Prices: Data vs Model

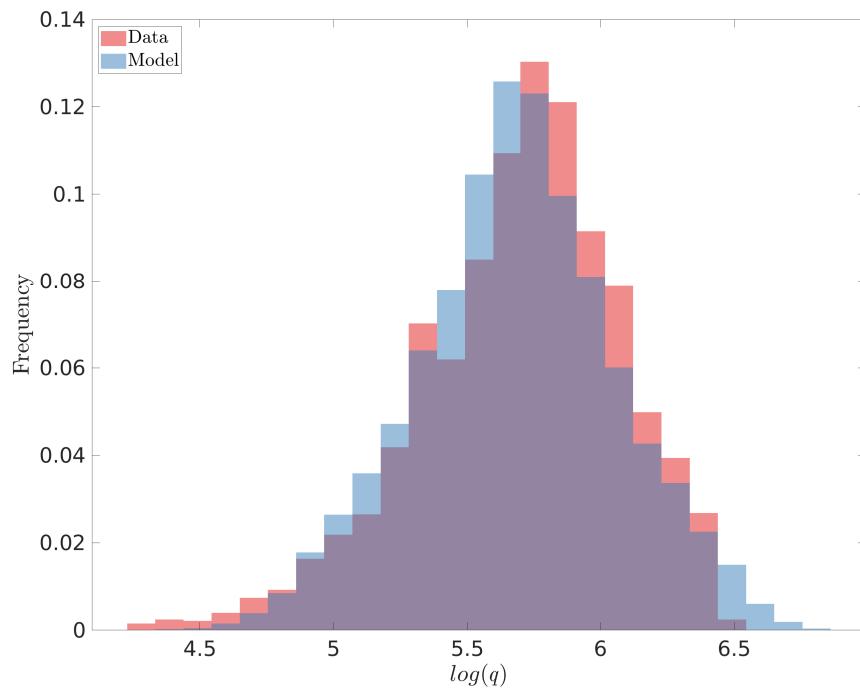


FIGURE 13: Histogram of Predicted Prices: Data vs Model

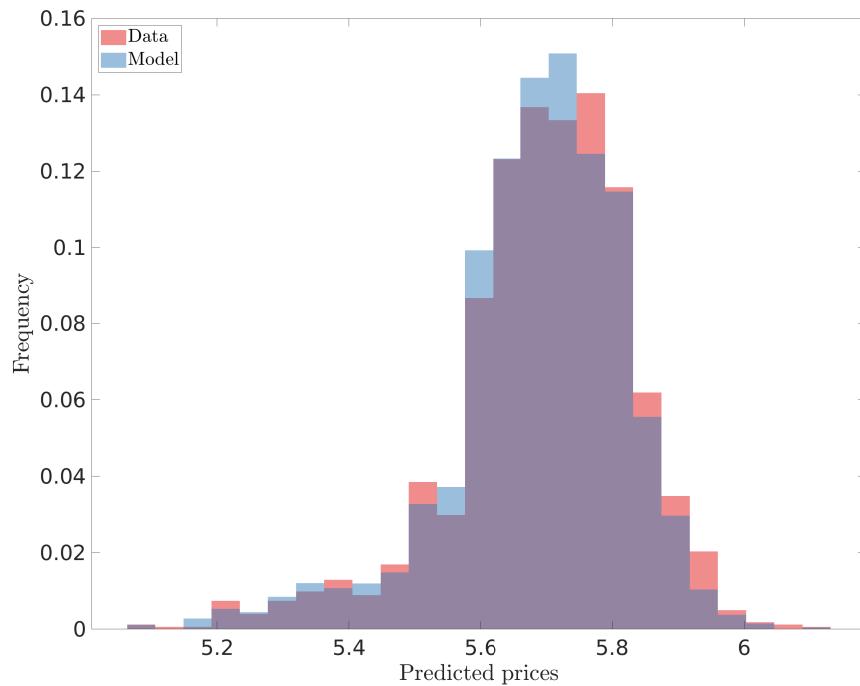


FIGURE 14: Histogram of Residual Prices: Data vs Model

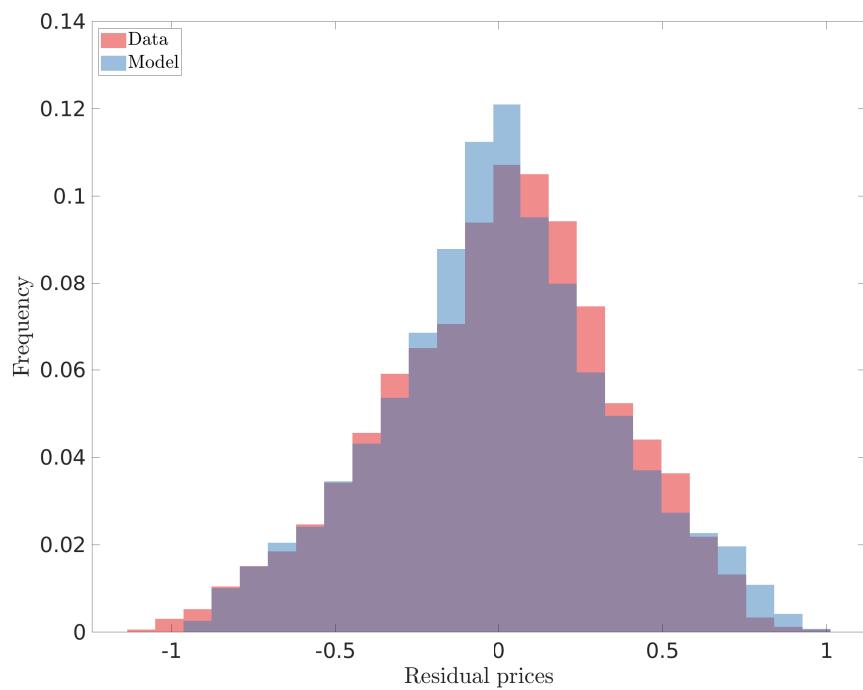


FIGURE 15: Histogram of Duration: Data vs Model

