Acquiring Innovation Under Information Frictions*

Murat Alp Celik
University of Toronto

Xu Tian
University of Georgia

Wenyu Wang
Indiana University

August 2021

Abstract

Acquiring innovation through M&A is subject to information frictions, as acquirers find it challenging to assess the value of innovative targets. We find an inverted U-shaped relation between firm innovation and takeover exposure; equity usage increases with target innovation; and deal completion rate drops with innovation. We develop and estimate a model of acquiring innovation under information frictions, featuring endogenous merger, innovation, and offer composition decisions. Our estimates suggest that acquirers' due diligence reveals only 30% of private information possessed by targets. Eliminating information frictions increases capitalized merger gains by 59%, stimulates innovation, and boosts productivity, business dynamism, and social welfare.

Keywords: information frictions, adverse selection, innovation, M&A.
JEL Classification: E20, G30, O40.

*Contact Information: Murat Alp Celik, 229 Max Gluskin House, 150 St. George Street, Department of Economics, University of Toronto, Toronto, Ontario, M5S 3G7, Canada. Email: murat.celik@utoronto.ca; Xu Tian, B340 Amos Hall, 620 South Lumpkin Street, Department of Finance, Terry College of Business, University of Georgia, Athens, GA 30602, United States. Email: xu.tian@uga.edu; Wenyu Wang, 1309 East Tenth Street, Department of Finance, Kelley School of Business, Indiana University, Bloomington, IN 47405, United States. Email: wenywang@indiana.edu. We would like to thank George Alessandria, Yan Bai, Andrey Golubov, Narayana Kocherlakota, Oliver Levine, Serdar Ozkan, Sergio Rebelo, Diego Restuccia, Xinxin Wang, Xiaodong Zhu, and seminar and conference participants at Indiana University, University of Toronto, Rochester Stockman Conference, SFS Cavalcade North America 2020, and SED Annual Meeting 2021 for their helpful comments and discussions. Wenyu Wang gratefully acknowledges financial support from the Daniel C. Smith Faculty Fellowship. All remaining errors are our own.

"Companies are buying innovation. As large companies need to be competitive and want to increase their footprints in a variety of different areas, one of the best ways to do that is through acquisition."

Peter Levine and Andreessen Horowitz

1 Introduction

Innovation is the key driver of economic growth. Recent literature documents that acquiring innovation and growth options is an important motive of corporate takeovers (e.g., Phillips and Zhdanov 2013, Bena and Li 2014, Levine 2017, and Wang 2018). The existence of an active M&A market, in which innovation can be efficiently reallocated, incentivizes many firms to specialize in innovation with the anticipation of being acquired later on. These firms create significant value for their shareholders by exiting with high offer premiums. They also stimulate the growth of the whole economy by reallocating their innovation to more efficient users. Reallocating innovation through M&A, however, is subject to significant information frictions. Outsiders often find it challenging to assess the value and the impact of corporate innovation, and the adverse selection risk can arise as a major concern for acquirers who seek to purchase innovative targets. Despite the substantial, yet largely separate literatures on information asymmetry, mergers and acquisitions, and corporate innovation, little has been done to understand the effect of information frictions on the reallocation of innovation through M&A. We aim to fill this gap in this paper.

We first document several novel findings regarding acquisitions of innovation. Using various measures that capture different aspects of corporate innovation, we document a robust inverted U-shaped relation between a firm's probability of being acquired and its stock of innovation. We also find that acquirers tend to use more equity in purchasing more innovative targets, and the probability of deal completion decreases with the target's innovativeness. These findings provide suggestive evidence for the role of information frictions in acquisitions of innovation. Specifically, as a firm becomes more innovative, its potential value as a target increases, which increases its exposure to takeovers. But meanwhile, there are a couple of counteracting forces at play. First, a firm's stand-alone value is also increasing in its own innovation, which lifts the bar for the firm to accept an offer. Second, assessing the value of an innovative target becomes increasingly challenging when the target possesses more private information, and the adverse selection risk faced by potential acquirers makes them less willing to pay an attractive price, which hinders the chance of striking a successful deal. Taken together, these opposing effects contribute to an inverted U-shaped relation between a firm's exposure to takeovers and its innovativeness.

Using equity as payment can help mitigate the adverse selection risk, because bringing the target shareholders on board forces them to share the risk with acquirer shareholders (e.g., Hansen 1987, Fishman 1989, and Eckbo, Giammarino, and Heinkel 1990). Acquirers, therefore, prefer using equity to purchase more innovative targets for hedging against the adverse selection risk. Equity usage, however, can also be costly to acquirers, because it is often accompanied by unfavorable market reactions to acquirer valuation due to the market's concern of the market timing hypothesis (e.g., Rhodes-Kropf, Robinson, and Viswanathan 2005, Rhodes-Kropf and Viswanathan 2004, Bhagat, Dong, Hirshleifer, and Noah 2005). The trade-off regarding the method of payment and the acquirers' reluctance of paying an attractive price, both driven by information frictions, reduce the deal completion rate when more innovative targets are involved.

Quantifying the effect of information frictions on the efficiency of innovation reallocation is challenging. We do not observe the acquirer's and the target's information sets, and therefore we cannot measure the adverse selection risk directly. More importantly, the M&A transactions and payment methods observed in the data are outcomes of an equilibrium in which acquirers and targets act strategically. To assess the effect of information frictions, we need to observe what would have happened in a parallel, counterfactual world in which information is perfect between the acquirers and targets. Measuring this counterfactual is difficult, because it is hard to find exogenous shocks that eliminate information frictions. Even if there were such a shock, it is likely to be limited in scope, raising concerns about external validity. Overall, it is unclear how to quantify the effect of information frictions without a model.

We overcome these challenges by estimating a model of acquiring innovation under information frictions. In the model, acquirers can create value by purchasing innovative targets. However, they sometimes cannot perfectly observe the target firms' stock of innovation. Instead, they form rational expectations over the true innovativeness based on the observables, and they make a take-it-or-leave-it offer to the targets. In the model, an acquirer has to decide on the amount of cash and equity to use in the offer, taking into account that the equity payment might be discounted. Upon receiving the offer, the target firm has to decide whether to accept it or not depending on its own information set and its continuation value. Firms in our model endogenously choose to become acquirers, targets, or remain stand-alone each period. In equilibrium, information frictions reduce acquirers' incentive to take risk and acquire targets with top-notch innovation, even though these targets can create the most value in M&A.

The model imposes no priors on whether information frictions are large, small, or even absent. We let the data tell us how important they are in transactions of innovation reallocation through M&A. We do so by estimating the model's parameters using the simulated method of moments (SMM). A few moments in the data are particularly important for our estimation. First, the inverted U-shaped relation between a firm's probability of being acquired and its stock of innovation, which we document in the data, reflects the two counteracting effects of innovation on firms' takeover exposure when information frictions are present. Second, acquirers endogenously choose the payment method by trading off the benefits of hedging against target value risk and the costs of equity payment being accepted at a discounted price. These patterns, therefore, help the model recover the extent of adverse selection risk faced by acquirers. The overall merger probability and the combined firm's announcement returns are informative of merger gains, which in turn discipline our estimates of the merger technology.

Our estimates suggest that the information asymmetry between acquirers and targets is substantial. Specifically, we estimate that, on average, due diligence conducted by acquirers helps reveal only 30% of private information possessed by the targets, and thus, acquirers face severe adverse selection risk in purchasing innovative targets. Information frictions, therefore, create a substantial barrier to trades between acquirers and targets.

Using the estimated model as a laboratory, we perform several counterfactual analyses to gauge the effect of information frictions on firm innovation and asset reallocation. We find that eliminating asymmetric information between the acquirers and targets has a strong, direct effect on value creation in the M&A market. As we decompose a firm's market value into its standalone value and the capitalized gains from M&A, we find that, with the estimated information frictions, the capitalized gains from the M&A market accounts for 6% of an average firm's market value, and eliminating information frictions in M&A boosts this ratio to 9.2%, representing an almost 60% increase. It also speeds up the process of asset reallocation such that innovative firms are bought much earlier in their life cycle and valuable innovations are reallocated to better users at a much faster pace.

A more efficient M&A market also triggers long-term effects on firm innovation and overall productivity growth: as asset reallocation becomes more efficient, firms optimally choose to invest more on R&D and conduct more innovation, especially small firms who aim to be bought in the future. The largest firms, on the other hand, adopt a lower innovation intensity, as they have the option of improving their productivity by acquiring productive small firms rather than spending too much on in-house R&D. Overall, the average innovation investment and firm growth rate increase in the economy.

We also quantify the value of equity payment in helping mitigate the effect of information frictions. When we prevent firms from using equity as payment, we find that the capitalized gains from the M&A market decline from 6% of firm value to less than 4%, representing a 38% decrease. This finding suggests that, despite its cost, equity payment remains an important vehicle that acquirers use to mitigate the adverse selection risk arising from purchasing innovative targets.

Our paper contributes to three strands of literature. First, it adds to the growing literature that investigates the role of reallocating corporate innovation in mergers and acquisitions. For example, Sevilir and Tian (2011) and Bena and Li (2014) find that combining complementary innovation capabilities creates significant synergies in M&A; Higgins and Rodriguez (2006) and Phillips and Zhdanov (2013) show that large and mature firms often optimally outsource their research and development by acquiring small and innovative firms; Wang (2018) documents that entrepreneurs cater to potential acquirers' demand in innovation. The two papers most related to our work are

Levine (2017) and David (2021). Levine (2017) builds a model in which acquirers pursue takeovers to obtain growth options from target firms. He shows that reallocation of growth options is another important motive of M&A beyond the traditional Q-theory of physical capital reallocation. David (2021) develops a search and matching model of M&A in general equilibrium, and uses it to evaluate the implications of merger activity for aggregate economic outcomes. It is found that the efficiency of the M&A market contributes significantly to the level of aggregate output and consumption. Our framework coincides with David (2021) in the synergistic merger technology and the use of a general equilibrium setting to assess the macroeconomic effects of the M&A market. The primary differences of our model are the endogenous productivity growth due to firm innovation, information frictions in the assessment of innovation that lead to adverse selection risk, and the use of equity vs. cash in merger transactions to reduce this risk. Recent work by Cunningham, Ederer, and Ma (2019) documents that some firms acquire innovative targets to preempt future competition (i.e., killer acquisition), and their estimate suggests that about 6.4% of acquisitions in their sample are killer acquisitions. Our paper documents a few novel findings that help identify information asymmetry as an important friction in acquiring innovation, and we quantify the effects of such information frictions in the economy.

Our paper also pertains to the literature on information asymmetry and the method of payment in M&A. Theoretical work by Hansen (1987), Fishman (1989), and Eckbo, Giammarino, and Heinkel (1990) model the payment method as a vehicle to signal acquirers' valuation of targets and to hedge the adverse selection risk caused by target private information. More recently, a few studies show that overvalued acquirers are more likely to use equity as cheap currency to purchase real assets from targets, and this market timing hypothesis suggests that equity payment is often accepted at a discounted price.[1] Vladimirov (2015) further argues that how cash bids are financed is equally important as the method of payment, because when acquirers lack access to competitive financing, they tend to finance their cash bids with equity and that often leads to underbidding and lower takeover premiums. Our model incorporates the method of payment as acquirers' endogenous

[1] See e.g., Shleifer and Vishny (2003), Rhodes-Kropf and Viswanathan (2004), Rhodes-Kropf, Robinson, and Viswanathan (2005), Bhagat, Dong, Hirshleifer, and Noah (2005), and Li, Taylor, and Wang (2018).

choice, and our estimates show that equity usage helps mitigate information frictions in acquiring innovation, but meanwhile it cannot fully insure acquirers from the adverse selection risk due to its cost to acquirers.

Methodologically, our paper belongs to the growing literature that employs structural model calibration or estimation to answer standard corporate finance questions in capital investment, leverage choice, CEO turnover, market competition and valuation, and particularly M&As, as summarized in Strebulaev and Whited (2012). In the field of M&As, Warusawitharana (2008) and Yang (2008) construct and estimate models that link asset purchases and sales to firms' productivity, and their findings are consistent with the neoclassical Q theory of M&As. Gorbenko and Malenko (2014) estimate valuations of strategic and financial bidders using a hand-collected data set of pre-announcement bids and conclude that different targets appeal to different types of bidders. Albuquerque and Schroth (2010, 2015) estimate structural models to quantify the private benefits of control in block trades. Eisfeldt and Rampini (2008) develop a dynamic model to study the effect of agency problem on capital reallocation over the business cycle. Dimopoulos and Sacchetto (2017) develop and calibrate a dynamic industry-equilibrium model to quantify the impact of merger activity on productive efficiency. Our paper contributes to this literature by estimating the effect of information frictions in acquisitions of innovation.

The paper is organized as follows. In Section 2, we present suggestive evidence to show that information frictions have a significant effect on innovation reallocation, which motivates our model. In Section 3, we introduce the model setup, and in Section 4, we present the model solution and demonstrate the main mechanisms. We discuss how we estimate the model and present our estimation results in Section 5. In Section 6, the estimated model is used to evaluate the effects of information frictions on innovation reallocation. Section 7 presents subsample estimation and robustness checks, and Section 8 concludes.

2 Suggestive Evidence

To motivate the pivotal role of information frictions in our model and estimation, in this section, we present a few empirical findings regarding the relation between mergers and acquisitions and firm innovation. They provide suggestive evidence on how information frictions affect firms' involvement in the M&A market and their M&A outcomes.

We first examine the relation between a firm's probability of being acquired and its innovation stock. We gauge a firm's innovation using several measures that capture different aspects of innovation, including the total number of patents, patent citations, patent originality, and breakthrough innovation (i.e., tail innovations). Table 1 summarizes the variable definitions.

We start with the univariate analysis in which we simply plot a firm's probability of being acquired against its innovation without any controls. Figure 1 shows the results for different innovation measures. There exists a robust hump-shaped relation between a firm's probability of being acquired and its innovation across all measures.

We further confirm our findings using multivariate regressions in which we control for a wide range of variables that are documented in previous studies to affect a firm's takeover exposure. Specifically, we perform the following regression analysis:

$$I^{acquired}_{i,t+1} = A_1 + B_1 \cdot Innov_{i,t} + C_1 \cdot Innov^2_{i,t} + \varrho_1 \cdot X_{i,t} + \varepsilon_{i,t} \tag{1}$$

where $I^{acquired}_{i,t+1}$ is a dummy which equals one if a firm i is acquired in year $t+1$, $Innov_{i,t}$ is one of the innovation measures for firm i in year t, $X_{i,t}$ is a vector of control variables including firm size, leverage, market-to-book ratio, return on assets, cash holding, R&D expenditure, asset tangibility, dividend payment dummy, firm age, the square of firm age, year fixed effects, and industry fixed effects. We include both $Innov_{i,t}$ and its quadratic term $Innov^2_{i,t}$ to capture the inverted U-shaped relation between a firm's probability of being acquired and its innovation. Table 2 presents the results. We find that even after controlling for a wide range of firm characteristics, the inverted U-shaped relation is still robust in the sense that for all innovation measures we use, we find a positive coefficient on the innovation measure and a negative coefficient on its quadratic term, both of which are statistically significant at 1% level in most regression specifications. The turning point, calculated as $innov^* = -\frac{B_1}{2C_1}$, also lies around the 50th to 75th percentile of each innovation measure, indicating that the inverted U-shaped relation is empirically relevant.[2] Taken together, we find that firms with an intermediate level of innovation are the most likely to be acquired, while firms with very weak or very strong innovation tend to stand alone. The inverted U-shaped relation is consistent with two counteracting forces brought about by firm innovation. On the one hand, innovative firms are more attractive targets in the M&A market, because they have the potential to create more value; on the other hand, if acquirers have imperfect information regarding target firm innovation, the adverse selection risk induced by information frictions makes it quite challenging to acquire highly innovative firms.

We next investigate the payment method in acquisitions of innovation. Previous studies show that the method of payment can also be influenced by information asymmetry (see e.g., Hansen 1987, Fishman 1989, and Eckbo, Giammarino, and Heinkel 1990). Specifically, if an acquirer is more concerned with information asymmetry about target quality, it may prefer using equity to reduce the adverse selection risk. This is because using equity brings the target shareholders on board and therefore any losses arising from adverse selection are shared between the acquirer shareholders and target shareholders. Instead, cash bids allow the target shareholders to cash out and the acquirer shareholders have to shoulder all adverse selection costs. We perform the following regression analysis regarding the payment method:

$$PrcOfStk_{i,t+1} = A_2 + B_2 \cdot Innov_{i,t} + \varrho_2 \cdot X_{i,t} + \varepsilon_{i,t} \tag{2}$$

where $PrcOfStk_{i,t+1}$ is the percentage of equity usage in a bid, $Innov_{i,t}$ is the target's innovation, and $X_{i,t}$ is a vector of control variables including acquirer innovation, acquirer and target misvaluation measures created by Rhodes-Kropf, Robinson, and Viswanathan (2005), acquirer and target size,

[2] We also perform a more rigorous hypothesis test for the existence of an inverted-U, following Lind and Mehlum (2010). The results are displayed in Table E4 in the Online Appendix, and they confirm the presence of an inverted-U relation between takeover exposure and our measures of innovation. See more details in Section D of the Online Appendix.

market-to-book ratio, return on assets, leverage, dummy for diversification merger, year fixed effects, and industry fixed effects. The results are reported in Table 3.

We find that equity usage indeed increases with the target's innovation in general, consistent with the argument that acquirers prefer using equity as payment to mitigate the adverse selection risk caused by information asymmetry regarding the target's innovation quality. We also find that equity usage increases with the acquirer's overvaluation, consistent with the market timing hypothesis. This analysis, therefore, highlights the potential costs and benefits of equity usage. On the one hand, when acquirers are uncertain about the targets' innovation quality, they prefer using equity as payment to hedge the adverse selection risk. On the other hand, equity usage may trigger investors' concern of market timing which in turn may lead to unfavorable market reaction to acquirers' valuation. These two counteracting forces determine the optimal choice of offer composition in merger transactions.

Last, we look at how the deal completion rate varies with the target's innovativeness. Since equity usage is costly and cannot eliminate the adverse selection risk, deal completion rate is expected to drop as information frictions intensify. Specifically, we perform the following regression analysis:

$$I_{i,t+1}^{complt} = A_3 + B_3 \cdot Innov_{i,t} + \varrho_3 \cdot X_{i,t} + \varepsilon_{i,t} \tag{3}$$

where $I_{i,t+1}^{complt}$ is the dummy that the deal completes, conditioning on firm i receiving a bid. The independent variables are defined as in regression Equation 2. In Table 4, we find that the probability of deal completion significantly decreases with the target's innovation across all four measures.

3 Model Setup

Motivated by the empirical evidence presented above, we now construct a model of acquiring innovation with information frictions. The model endogenizes firms' innovation and M&A decisions, and the method of payment used in these transactions.

In the model, time is discrete and denoted by $t \in \{0, 1, 2, ...\}$. There is a continuum of firms, indexed by $i \in [0,1]$, which are heterogeneous in productivity $z_{it} > 0$. Firms use capital to produce a homogeneous final consumption good. We specify the production function in Section 3.2 below. As we will show later, decreasing returns to scale implies that a firm's flow profit is increasing in its productivity z_{it}. Firms invest in costly R&D to maintain and improve their productivity z_{it} over time.

Firms can also participate in a merger market where they act as an acquirer or a target. The M&A market features random matching: each period, a firm meets with a potential merging partner, and they decide whether to merge together. We assume that firms' contemporaneous productivity is privately observed, but conditional on a matching, with some probability, the acquirer firm can observe the contemporaneous productivity of the target through due diligence. This probability captures the information friction in our model. If the contemporaneous productivity is not revealed, the acquirer forms a rational expectation based on the target's productivity in previous periods.

The acquirer then makes a take-it-or-leave-it offer to the target firm and specifies the method of payment (cash, equity, or a mixed offer). After receiving the offer, the target firm decides whether to accept it or not. This decision hinges on the future opportunities available to the target as well as the expected offer value perceived by the target. If the merger completes, two firms merge into one combined firm, and the target firm ceases to exist. We also assume an exogenous rate of exit of all firms in the model. All exiting firms (due to endogenous acquisition or exogenous exit) are replaced by new firms such that the total measure of firms remains constant.

3.1 Household

We assume a representative household with time-separable preferences exhibiting constant relative risk aversion. The time discount rate is $\beta \in (0,1)$. The solution to the standard household optimization problem delivers the Euler equation $r = \beta^{-1} - 1$ in a stationary equilibrium, which governs the risk-free discount rate, r, in our model.

3.2 Firm Production and Innovation

Each period, firms produce a homogeneous final consumption good in our model with the Cobb-Douglas production function:

$$y_{it} = z_{it}^{\zeta} k_{it}^{\kappa} \tag{4}$$

where z_{it} is the firm-specific productivity, and k_{it} is the capital input. We impose $\zeta \in (0,1)$, $\kappa \in (0,1)$, and $\zeta + \kappa = 1$. Since $\kappa < 1$, the production function features decreasing returns to scale.

We assume that there is no adjustment cost of capital, and the capital depreciation rate is $\delta \in (0,1)$. A firm chooses optimal capital, within each period, to maximize its period profits:

$$\Pi_t(z_{it}) = \max_{k_{it} \geq 0} \left\{ z_{it}^{\zeta} k_{it}^{\kappa} - (r+\delta) k_{it} \right\} \tag{5}$$

Solving the profit maximization problem above yields the firm's flow profit as follows:

$$\Pi_t(z_{it}) = \pi z_{it} \tag{6}$$

where $\pi \equiv \zeta \left(\frac{\kappa}{r+\delta}\right)^{\frac{\kappa}{1-\kappa}}$ is a time-invariant constant in a stationary equilibrium. The firm's period profits are therefore linear in its productivity.

Firm productivity z_{it} fluctuates over time, and firms innovate to maintain and improve their productivity. Specifically, we model firm productivity as an AR(1) process:

$$\log z_{it+1} = \mu_{it} + \rho \log z_{it} + \sigma_{\epsilon} \epsilon_{it+1} \tag{7}$$

We further assume that firm innovation affects μ_{it}, which influences the long-run mean of productivity. This specification captures the idea that innovation creates growth options for generating future projects and it pushes forward a firm's technology frontier (e.g., Levine 2017). In order to maintain a long-run mean of productivity μ_{it}, firms incur research and development costs of innovation each period:

$$X(\mu_{it}, z_{it}) = \chi(\mu_{it} - \underline{\mu})^2 z_{it} \tag{8}$$

where $\chi > 0$ is a scale parameter, and $\underline{\mu}$ is the lower bound attained when the firm chooses not to innovate (i.e., zero expenditure on R&D). The linear term z_{it} ensures that the cost of R&D scales up with the gains from innovation, a standard feature in models with endogenous productivity growth.

Firms choose their optimal innovation policy (i.e., the optimal μ_{it}) each period by maximizing the expected gains from innovation net of the R&D cost. The value of innovation arises from two sources: first, innovation improves a firm's productivity and thus its stand-alone value; and second, a more innovative firm demands a higher offer price when acquired and thus provides a better value at exit. We characterize firm-optimal innovation decisions in Section 3.4 after introducing the market for M&A.

3.3 The Market for M&A

3.3.1 Merger Technology

Firms also participate in an active M&A market and can potentially benefit from merging with other firms. Our model features a random matching between firms. Specifically, each period, a firm meets with a potential merging partner, and upon the match, two firms examine the expected gains from merging together, which depend on the combined firm's productivity. We assume a Cobb-Douglas merger technology as in David (2021), that is, as an acquirer with productivity z_A merges with a target with productivity z_T, the combined firm's productivity $z_m = f(z_A, z_T)$ is determined as:

$$f(z_A, z_T) = \gamma z_A^{\alpha} z_T^{1-\alpha} \tag{9}$$

with $\gamma > 0$, and $\alpha \in (0, 1)$. The parameter $\gamma > 0$ controls the scale of total value created in the merger, and the parameter α controls the relative share of contribution to the combined firm's productivity made by the acquirer and target. Intuitively, a higher value of $\gamma > 0$ implies more profitable mergers, and a higher value of α amplifies the role of the acquirer in the merger.

Equation 9 suggests that the combined firm's productivity depends on the role of the two merging firms in the M&A and therefore our model also characterizes the endogenous decision of "who acquires whom" in equilibrium.

3.3.2 Information Structure in M&A

Information asymmetry is the key friction in our model, and it affects both merger terms (e.g., offer value, method of payment) and outcomes. We build this friction into our model by assuming asymmetric information on firm productivity z. Specifically, we assume that each firm i perfectly observes its own productivity at any point of time.[3] Other firms, however, can only observe firm i's productivity with a one-period lag. This is because any firm's past profits are observable and Equation 6 suggests that productivity is linear in profits. As a result, public information of lagged profits constitutes the first channel of information flow, through which corporate outsiders learn about firm private information.

Mergers create the second channel of information flow in the model. As two firms meet and contemplate a potential merger, they spend more time learning about each other. The acquirer conducts due diligence before making an offer, and we assume that, with probability $1 - \omega \in (0, 1)$, the target's contemporaneous productivity is fully revealed, and thus the acquirer's information set becomes the same as the target's; with the complementary probability ω, the target's contemporaneous productivity remains private, and in this case, the acquirer has to form a rational expectation based on the target's lagged productivity. A larger value of ω therefore is associated with more substantial information frictions faced by the acquirer.

For the target firm, we assume that it always learns the acquirer's contemporaneous productivity through contact. Arguably, the main challenge of acquiring innovation results from the difficulty in evaluating target innovation, and thus the target-side, rather than the acquirer-side, private

[3] We relax this assumption in Section 7.2, where we consider the possibility of firms not knowing their own contemporaneous productivity either.

information plays the pivotal role in our model. We discuss this assumption in more detail in Section 3.5 and analyze the effect of the acquirer-side information asymmetry in Online Appendix B.1.

3.3.3 Timeline of Merger

We characterize a typical merger process in Figure 2. For a given period t, the figure illustrates the series of events (in blue) a firm experiences and the sequence of actions (in red) it takes during this process.

Specifically, at the beginning of period t, the firm chooses its innovation level, μ, which affects its contemporaneous productivity, governed by Equation 7. The firm then randomly matches with a potential merger partner. Upon the match, the firm has to choose its potential role in the merger, either as the acquirer or as the target, and this decision generates two branches on the timeline.

If the firm acts as the target (the upper branch path), it learns the contemporaneous productivity of itself and that of the acquirer (i.e., z and z_A) immediately after the innovation outcomes are realized. The acquirer then conducts due diligence. With probability $1 - \omega$, the firm's contemporaneous productivity, z, is revealed to the acquirer, and with probability ω, the firm's contemporaneous productivity remains its private information. The firm then receives an offer from the acquirer that specifies the method of payment. The offer depends on the acquirer's information set after due diligence (i.e., whether the firm's contemporaneous productivity is revealed). The firm then decides whether to accept the offer and merge with the acquirer or to decline the offer and remain standalone.

If the firm acts as the acquirer (the lower branch path), it learns its own contemporaneous productivity immediately after the innovation outcomes are realized. It, however, does not observe the target's new productivity z_T right away. The firm then conducts due diligence that reveals the target's z_T with a probability of $1 - \omega$. Based on the firm's information set after due diligence, it makes an offer to the target, which may be accepted or declined. If an offer is accepted, then two matched firms merge together and the target firm exits; if an offer is declined, then both firms remain standalone.

We summarize the timeline of the merger process as follows:

1. Firms choose their innovation policy μ_i and incur the R&D cost $X(\mu_i, z)$.

2. Matches in the merger market are realized.

3. Two matched firms choose their roles as the acquirer or as the target.

4. Innovation outcomes are realized, and the contemporaneous productivity of the acquirer and the target are updated.

5. With probability $1-\omega$, the target's contemporaneous productivity is revealed to the acquirer, and with probability ω, it remains the private information of the target.

6. The acquirer firm makes an offer to the target, specifying the method of payment.

7. The target decides whether to accept or decline the offer.

3.3.4 Decision to Merge

Acquiring firm: First, we consider the problem of a given acquiring firm with technology z_A that meets a target firm. There are two possible scenarios:

<u>Scenario 1.</u> If the target's contemporaneous productivity is not revealed, the acquirer only observes the target's last period productivity $\tilde{z}_T$ and would try to form a rational expectation of the target's contemporaneous productivity z_T based on its information set. We denote the value function of the acquirer under this scenario as $V_A^i(z_A, \tilde{z}_T; \Theta)$, where the superscript i stands for *imperfect information*. The maximization problem is set up as:

$$\begin{aligned} V_A^i(z_A, \tilde{z}_T; \Theta) = \max_{c \geq 0, s \in [0,1]} \quad \Bigg\{ \quad & \mathbb{E}\Bigg[\mathbb{I}_T \Bigg((1-s) \Bigg(-c + \pi z_m - g(c, s, z_m; \Theta') + \frac{1-\psi}{1+r} V(z_m; \Theta') \Bigg) \Bigg) \\ & + (1-\mathbb{I}_T) \left(\pi z_A + \frac{1-\psi}{1+r} V(z_A; \Theta') \right) \Bigg| \tilde{z}_T, z_A; \Theta \Bigg] \Bigg\} \qquad (10) \end{aligned}$$

$$\mathbb{I}_T = \hat{\mathbb{I}}_T(z_A, z_T, c, s; \Theta) \qquad (11)$$

where Θ denotes the aggregate state of the economy, r is the risk-free discount rate, and ψ is the probability of exogenous exit.[4] The acquirer chooses the optimal amount of cash c and the fraction of the combined firm's equity s to offer to the target. The first term in the above equation denotes the value to the acquirer if the offer is accepted ($\mathbb{I}_T = 1$), and the second term denotes the value to the acquirer if the offer is rejected.[5] If the offer is accepted, the acquirer pays c to the target and gets $1 - s$ fraction of the combined firm value. Empirically, equity payment is often accompanied by unfavorable market reactions to the acquirer's valuation, which constitute the costs of equity usage. Motivated by this empirical fact, we deduct a term $g(c, s, z_m; \Theta')$ from the combined firm's value:

$$g(c, s, z_m; \Theta') = \Lambda E(c, s, z_m; \Theta') z_m \tag{12}$$

where the term $E(c, s, z_m; \Theta')$ is the equity ratio of the offer. With $\Lambda > 0$, the expected value loss due to equity usage is increasing in the equity ratio, and the costs of equity usage are zero if the offer is paid with 100% cash (i.e., $s = 0$).

We choose to capture the costs of equity usage with the function $g(c, s, z_m; \Theta')$ for two reasons. First, our model characterizes the optimal offer choices made by an acquirer in face of the costs of equity usage, and therefore the magnitude of the costs is crucial to the model while its microfoundation is of less importance. In other words, as long as the costs of equity usage are properly specified, the model solution and estimation are robust to the sources of such costs. Second, as we will discuss in Section 3.5, the specification of function $g(c, s, z_m; \Theta')$ is consistent with an extended model that incorporates the acquirer-side information asymmetry. Such an extension, however, does not provide additional insights to the model beyond the reduced-form specification of the function $g(c, s, z_m; \Theta')$, but would significantly complicate the model, as well as lead to the problem of equilibrium selection.

We denote the cash and equity components of the optimal offer, solved endogenously in the model, as $\hat{c}^i(z_A, \tilde{z}_T; \Theta)$ and $\hat{s}^i(z_A, \tilde{z}_T; \Theta)$ when the target's true productivity is not known. The hat notation is used to indicate that these functions are the optimal policy rules. In the rest of the paper,

[4] The aggregate state deterministically evolves to Θ', and in a stationary equilibrium, $\Theta' = \Theta, \forall t$.

[5] This is determined by the target's maximization problem we describe below shortly. In equilibrium, rational expectation fulfills so that the acquirer correctly anticipates the target's decision rule denoted by $\hat{\mathbb{I}}_T(z_A, z_T, c, s; \Theta)$.

we continue to use the hat notation to indicate optimal policy rules.

Scenario 2. If the target's contemporaneous productivity is fully revealed, then the acquirer perfectly observes z_T, and the acquirer's maximization problem is simpler:

$$
\begin{aligned}
V_A^p(z_A, z_T; \Theta) = \max_{c \geq 0, s \in [0,1]} \quad \Bigg\{ \quad & \mathbb{I}_T \left((1-s) \left(-c + \pi z_m - g(c, s, z_m; \Theta') + \frac{1-\psi}{1+r} V(z_m; \Theta') \right) \right) \\
& + (1 - \mathbb{I}_T) \left(\pi z_A + \frac{1-\psi}{1+r} V(z_A; \Theta') \right) \Bigg\} \qquad (13)
\end{aligned}
$$

$$
\mathbb{I}_T \quad = \quad \hat{\mathbb{I}}_T(z_A, z_T, c, s; \Theta) \qquad (14)
$$

In this case, we denote the cash and equity components of the optimal offer as $\hat{c}^p(z_A, z_T; \Theta)$ and $\hat{s}^p(z_A, z_T; \Theta)$. We use the superscript p for the value function and the optimal policy rules to denote this scenario under which the acquirer has *perfect information*.

Target firm: Now, consider the problem of a target firm with technology z_T which receives a merger offer of cash and equity, (c, s), from an acquirer firm with technology z_A. The target firm decides whether to accept or reject the offer by comparing its continuation value and the offer price:

$$
\begin{aligned}
V_T(z_A, z_T, c, s; \Theta) = \max_{\mathbb{I}_T \in \{0,1\}} \quad \Bigg\{ \quad & \mathbb{I}_T \left[c + s \left(-c + \pi z_m - g(c, s, z_m; \Theta') + \frac{1-\psi}{1+r} V(z_m; \Theta') \right) \right] \\
& + (1 - \mathbb{I}_T) \left[\pi z_T + \frac{1-\psi}{1+r} V(z_T; \Theta') \right] \Bigg\} \qquad (15)
\end{aligned}
$$

The first term represents the offer value perceived by the target and the second term is the target's continuation value if it rejects the offer. We denote the acceptance/rejection decision that solves the above maximization problem as $\hat{\mathbb{I}}_T(z_A, z_T, c, s; \Theta)$ accordingly.

3.4 Firm Innovation Decision

Given the expected merger gains described above, we now can characterize firms' optimal innovation decisions. At the beginning of each period, firms choose their innovation policy μ_i to maximize the expected firm value:

$$
\begin{aligned}
V(z;\Theta) &= \max_{\mu_i \geq \underline{\mu}} \Bigg\{ \int \Bigg[\mathbb{I}_{acq}(z,z_o) \Big(\omega \mathbb{E} \left[V_A^i(z', z_o; \Theta) \Big| z, \mu_i; \Theta \right] \\
&+ (1-\omega) \mathbb{E} \left[V_A^p(z', z_o'; \Theta) \middle| z, \mu_i; \Theta \right] \Big) \\
&+ (1 - \mathbb{I}_{acq}(z,z_o)) \Big(\omega \mathbb{E} \left[V_T(z_o', z', \hat{c}^i(z_o', z), \hat{s}^i(z_o', z); \Theta) \Big| z, \mu_i; \Theta \right] \\
&+ (1-\omega) \mathbb{E} \left[V_T(z_o', z', \hat{c}^p(z_o', z'), \hat{s}^p(z_o', z'); \Theta) \middle| z, \mu_i; \Theta \right] \Big) \Bigg] dF_s(z_o) \\
&- X(\mu_i, z) \Bigg\}
\end{aligned} \tag{16}
$$

where $\mathbb{I}_{acq}(z, z_o)$ is an indicator function that takes the value one if the current firm acts as the acquirer in a given merger and zero otherwise,[6] and $F_s(z)$ is the steady-state distribution of firms at the beginning of the period. z_o denotes the matched firm's productivity. Since matching is random, $z_o \sim F_s(z)$. Because innovation realizations occur after the acquirer and the target firms are chosen, new productivites z' and z_o' are random variables that depend on the productivities at the beginning of the period z and z_o, and the associated innovation policies. The first term stands for the case when the firm is an acquirer and has imperfect information about the target's productivity z_o'. The second term is for the case when the firm is an acquirer and has perfect information regarding the target. The third term stands for the case where the firm is a target, and receives an offer from an acquirer acting under imperfect information. The fourth term is the case where the firm is once again a target, but the acquirer knows the firm's productivity z'. The final term is the cost of R&D necessary to maintain the innovation level μ_i. Denote the optimal innovation policy function that solves the problem as $\hat{\mu}_i(z;\Theta)$.

3.5 Model Discussion

Like any economic model, our model makes a few important simplifying assumptions that allow us to focus on the mechanisms of interest while maintaining tractability. In this section, we discuss

[6]The optimal decision to be the acquirer or the target maximizes ex-ante surplus from the merger.

their effects on our model solution and estimation.

First and foremost, our model focuses on characterizing the adverse selection risk that arises from the target-side private information. We believe that this focus is central to the problem we study in the paper. It is, however, also possible that the acquirers possess private information. In fact, the market timing hypothesis conjectures that acquirer-side private information induces overvalued acquirers to use equity as cheap currency in their takeover bids, which may trigger negative market reactions to equity usage. Since the payment method is crucial to the model solution, we capture the effect of acquirer-side private information on the method of payment by specifying a function $g(c, s, z_m; \Theta')$. This function characterizes the acquirer's costs of using equity in takeovers, and we let the data guide our estimation of the function parameters. In this sense, our estimation indeed "controls for" the effect of acquirer-side private information, and as long as the function $g(c, s, z_m; \Theta')$ is properly specified and estimated, our estimates are likely unbiased.

In Online Appendix B.1, we present an extended model that explicitly incorporates the acquirer-side private information. We show that such a two-sided information asymmetry model gives rise to the well-known equilibrium multiplicity outcome and equilibrium selection becomes necessary. An appealing feature, however, is that all these equilibria predict the same discounting of equity by the target, which is consistent with our specification of $g(c, s, z_m; \Theta')$. This extension therefore provides a microfoundation for the function $g(c, s, z_m; \Theta')$ in our baseline model.

Second, mergers and acquisitions can be driven by various motives. The merger technology we specify in Equation 9 nests the Q-theory of M&A (Jovanovic and Rousseau 2002) as α approaches 1 and the asset complementarity theory (Rhodes-Kropf and Robinson 2008) as α approaches 0.5. Our model allows firm innovation to affect firm productivity and thus naturally embeds the incentive of acquiring innovation. Our model, however, does not feature imperfect competition. Using a model of M&A with oligopolistic competition in each industry, Cavenaile, Celik, and Tian (2021) show that only mergers between superstar firms in two superstar industries are "anti-competitive" in nature, whereas the mergers in remaining industry states are motivated by synergy gains instead. Cunningham, Ederer, and Ma (2019) also find that only about 6% of acquisitions in pharmaceuticals industry are killer acquisitions. Given the relatively small fraction of acquisitions with anti-competitive implications, we expect that leaving this motive out of our model would not significantly bias our estimates in our sample.

Third, our model makes a parsimonious assumption that information frictions arise from the latest shocks to the target's innovation outcome. This specification is consistent with the model setup that firm profitability is proportional to past innovation outcomes and thus observing historical profitability reveals innovation outcomes in the past. Following the innovation literature, we present motivating empirical evidence (in Section 2) and estimate the model parameters (in Section 5) using innovation stock data in the paper. In Online Appendix E Table E1 to E3, we show that the empirical results that motivate our study and guide our estimation remain quantitatively similar if we use the innovation measured during the past 3 years instead, arguably a better proxy for firms' latest innovation activities.

Last but not least, the model assumes that target firms observe their own contemporaneous innovation realization before accepting/declining the M&A offer. One possible concern is whether this assumption is warranted and whether violation of this assumption would significantly bias our results. In Section 7.2, we extend the model to allow target firms to have imperfect information about their own innovation and evaluate the robustness of our results. We demonstrate that the inefficiencies caused by information frictions remain quite substantial even if target firms possess no private information in 30% of cases.

4 Model Solution

We solve the model numerically and characterize the model solution below. See Online Appendix A for computational details. The model solution describes the acquirer's decision on offer price, the target's decision on whether to accept or decline the offer, and the firm's innovation policies. We also discuss how merger probability and deal completion rate are determined in the equilibrium.

4.1 Acquirer Offer Price

Acquirers need to determine the offers made to the targets. The offer price depends on whether the target's quality is revealed or not. When the target's quality is fully revealed, there is no information asymmetry between the two firms and the offer price is easy to determine. Our analyses below, therefore, focus on the situation when the target's quality is not fully revealed and the acquirer needs to form a rational expectation of the target's quality. Our results show that there exist cases in which the acquirer cannot perfectly screen the target's type and it makes an offer that is insufficient to attract the target of best quality. In this sense, information asymmetry leads to inefficiencies in the takeover market because the deals that have the potential of creating the largest value fail to consummate.

To elucidate the mechanism, we consider an example where an acquirer with type z_A meets a target with pre-innovation productivity $\tilde{z}_T$. The true productivity of the target, z_T, is not fully revealed, so the acquirer must conjecture based on the innovation policy of the target, $\hat{\mu}_i(\tilde{z}_T;\Theta)$ and the law of motion for productivity given by Equation 7. For simplicity, let's consider a discretized version of the model with only three possible types for the target: $z_T^h > z_T^m > z_T^l$ with $z_T^m = \tilde{z}_T$.[7] The acquirer knows how each of the three types would react to an offer (c,s). Figure 3 depicts the indifference curves of each target type, along which they are indifferent between accepting or rejecting the offer. Any offer above the curve would be accepted. In this example, an offer that is acceptable to z_T^h would be acceptable to both z_T^m and z_T^l, but the opposite is not true. Given this information, the acquirer has to choose (c,s) to maximize its expected value.

Given the target's indifference curve that is perfectly anticipated by the acquirer, the left panel of Figure 4 presents a heat map that illustrates the acquirer's value by making different offers. Lighter color indicates higher value to the acquirer: the acquirer's value is low at the northeast corner where the acquirer makes a very generous offer and overpays the target to a large extent. The acquirer's value is also low at the southwest corner where the offer made by the acquirer is too low, and the deal fails to complete and thus generates no merger gains for the acquirer. The optimal offer price is

[7] That is, a Tauchen discretization of $\mathbb{E}[z_T|\tilde{z}_T, \hat{\mu}_i(\tilde{z}_T;\Theta)]$ with three bins.

set based on the acquirer's indifference curve that maximizes his expected gains, taking into account the target's response determined by the target indifference curve shown in Figure 3. In our example, the value function of the acquirer is maximized at the red dot, and the offer contains both cash and equity. This offer is accepted by z_T^l and z_T^m, but rejected by z_T^h.

The right panel of Figure 4 presents the heat map that illustrates the value function of a target with z_T^m. The value function is increasing in c and s above the indifference curve, because any offers with value above the indifference curve will be accepted and a more lucrative offer increases the target's value. However, the value function is flat below the indifference curve, because any offer below that will be rejected and then the value function is simply equal to the target's continuation value. Therefore, the target type that marginally accepts the offer, z_T^m, is only paid its outside option, whereas the low type z_T^l is paid in excess of its outside option and collects information rents. This logic carries over to cases with more target types or even continuous types.

4.2 Target Response

We now present the target's response to takeover offers. If the target type is fully revealed through the acquirer's due diligence, the acquirer will make an offer that is always acceptable to the target as long as the surplus from merger is positive. Our analyses here therefore focus on the situation when the target type is not fully revealed. As we discussed above, the acquirer may make an offer that is appealing to only a subset of targets in this case. Let's assume that the acquirer has a productivity z_A and makes the optimal offer as discussed in the above section. We still assume that there are three types of targets and denote their pre-innovation productivity $\tilde{z}_T$.

Figure 5 plots the deal acceptance policy functions $\hat{\mathbb{I}}_T(z_A, z_T, \hat{c}^i(z_A, \tilde{z}_T), \hat{s}^i(z_A, \tilde{z}_T))$ of the three post-innovation types defined as before for multiple values of $\tilde{z}_T$, where a value of one indicates acceptance.

When $\tilde{z}_T$ is low, the acquirer makes an offer that would be accepted by all types. This happens because, as Equation 7 suggests, the effect of innovation scales with the level of the lagged productivity $\tilde{z}_T$, and therefore when $\tilde{z}_T$ is low, uncertainty caused by the unobservable innovation is small in magnitude. In this case, the adverse selection problem induced by information frictions is less of a concern to the acquirer and it is able to make an offer that is generous enough to be accepted by all types of targets. In other words, the expected synergy is high enough to compensate the acquirer for the risk of overpaying the low and medium type of targets.

However, as $\tilde{z}_T$ increases, the adverse selection risk induced by information frictions intensifies and becomes more costly to the acquirer. The acquirer thus makes a more conservative offer, which is only accepted by the medium and low type targets. Even though the acquirer anticipates to lose high quality targets, it chooses to do so because the benefits of making a higher offer to attract high quality targets are overwhelmed by the costs of significantly overpaying the lower quality targets.

As $\tilde{z}_T$ becomes very high, the acquirer becomes even more conservative, because it understands that the costs of overpaying a low quality target can be quite substantial in this case. That is, if it makes a very lucrative offer but eventually ends up with acquiring a low type of target, the deal can come as a big disappointment. As a result, the acquirer chooses to make a very conservative offer that is accepted by only the low type target. This is the classical market for lemons problem studied in Akerlof (1970). As adverse selection risk aggravates, it leads to the separation of low types from the medium and high types instead of the pooling results we observed above.

4.3 Innovation Policy and Firm Growth

Figure 6 depicts the innovation policy $\hat{\mu}_i(z;\Theta)$, and the implied expected firm growth rate for firms with different productivity levels z in the left and right panels, respectively. Consistent with the empirical findings in previous literature that reject Gibrat's law (See Hall (1987) among others), smaller firms innovate more intensively and grow faster. The existence of an active M&A market creates an extra incentive for small firms to innovate, as becoming more productive makes them more valuable to potential acquirers. Large firms can choose a lower innovation intensity, as they have the option of improving their productivity by acquiring productive small firms rather than spending too much on in-house R&D.

4.4 Merger Probability, Deal Completion, and Offer Composition

The above sections present model solutions as functions of state variables and illustrate the economic intuition behind them. To better link our model to the data, we examine in this section how the model-implied merger probability, deal completion rate, and the method of payment vary with the target firm's simulated stock of innovation in model equilibrium.[8] To do so, we simulate the model and generate a cross-section of randomly matched pairs of potential acquirers and targets.[9] We track the deal outcomes and the method of payment in each deal. We then group the simulated data into bins based on target firms' stock of innovation and calculate the average merger probability, deal completion rate, and equity ratio within each bin. Figures 7 to 9 demonstrate our analyses.

Figure 7 depicts the model-implied relation between firm innovativeness and the probability of being acquired. When firm innovativeness is quite low, the benefit from M&A is small and thus the probability of being acquired is close to zero. As a firm becomes more innovative, its potential contribution as a target rises and thus the likelihood of being acquired gradually increases. The increased expected merger gain, however, is counteracted by the growing adverse selection risk, eventually leading to a hump-shaped relation between the takeover exposure and firm innovativeness. This happens because as a firm becomes more innovative, the realization of its ongoing innovation has a larger impact on its value and the expected synergy in M&A, leading to an intensified adverse selection risk. As a result, highly innovative firms are more difficult to be acquired due to the information friction.

Confirming our above discussion, Figure 8 illustrates a negative relation between the deal completion rate and target innovativeness. This happens because, facing an intensified adverse selection risk, acquirers tend to make more conservative offers to more innovative targets, and such offers are more likely to be turned down by targets of high quality.

Last, Figure 9 shows that there is a positive relation between the equity ratio of the offer and target innovativeness. Since the acquirers cannot always perfectly observe the target's true

[8] We construct the discounted stock of innovation for the simulated firms by using the same methodology as in the empirical analysis. The details can be found in Online Appendix A.

[9] We use the estimated parameter values reported in Table 5 to solve and simulate the model.

productivity under information asymmetry, they hedge the adverse selection risk by using equity in their offers. Adverse selection risk increases with the target firm innovation, and this leads to an increased usage of equity as the targets become more innovative.

5 Estimation and Identification

5.1 Identification

We start with parameters that are standard in the literature. The first three parameters are the subjective discount rate β, capital depreciation rate δ, and the production function concavity on capital κ. We set $\beta = 0.9615$, consistent with a real interest rate of $r = 0.04$, $\delta = 0.069$ is taken from U.S. NIPA, and $\kappa = 0.85$, consistent with an average markup of 18%.

We are left with 9 parameters to be estimated: $\omega, \Lambda, \chi, \rho, \sigma, \underline{\mu}, \psi, \gamma$ and α. We use 14 moments to identify the remaining 9 model parameters. Our identification strategy ensures that there is a unique parameter vector that makes the model fit the data as closely as possible. Since we estimate these parameters in one big SMM system, we essentially allow each moment to respond to all parameters in estimation. We explain below which moments are the most informative in identifying each parameter.

1. **Information asymmetry:** The parameter ω determines the method of payment in our model. As the acquirer cannot always observe the target's contemporaneous productivity, it may prefer using equity for hedging purpose. Therefore, a higher value of ω increases the overall equity usage in the model. We thus use the average equity ratio of offers and the sensitivity of equity usage with respect to target innovativeness to help pin down ω. In particular, we repeat the regression outlined in Equation 2 using the model-simulated data, and seek to match the model-implied coefficient B_2 therein.

   Meanwhile, since equity usage is costly, acquirers are more reluctant to make a lucrative offer with equity payment, which reduces the likelihood of the offer being accepted. This is particularly true when acquirers bid for highly innovative targets where equity usage is crucial.

So we also include the sensitivity of deal completion rate with respect to the target's stock of innovation as a targeted moment in SMM. Namely, we run the regression given in Equation 3 using the model-simulated data, and match the coefficient B_3.

2. **Cost of equity usage:** Consistent with empirical evidence, using equity is costly for acquirers due to the negative market reaction. The magnitude of this cost is governed by Λ. We estimate the value of Λ by replicating the empirical relation between equity usage and announcement returns. This is done by targeting the implied value loss for acquirers obtained from the following regression:

$$AcqRet = A_4 + B_4 \cdot PrcOfStk + \varrho_4 \cdot X_{i,t} + \varepsilon_{i,t} \tag{17}$$

Acquirer announcement return is regressed on the equity ratio, and the controls $X_{i,t}$ include the relative size of target, the size of acquirer, leverage, market-to-book ratio, return on assets, and cash to assets of the acquirer and target, as well as year and industry fixed effects. The implied value loss for acquirers is calculated as $|B_4 \cdot PrcOfStk|$.

3. **R&D cost:** The scale parameter of the R&D cost function, χ, determines the amount of resources firms spend on innovation. We target average R&D intensity defined as R&D expenditure over total assets to determine its value – a standard moment used in endogenous growth models.

4. **Acquirer's share of contribution in M&A:** The productivity of a combined firm is a Cobb-Douglas function of the productivities of the acquirer and the target. The parameter α determines the relative contribution of the acquiring firm. The relative size of the target firm to the acquirer helps identify α. As α increases, the acquiring firm contributes a larger fraction to the combined firm and thus the ratio of target-to-acquirer size shrinks accordingly.

5. **Evolution of firm productivity:** A firm's productivity evolves according to the AR(1) process given in Equation 7. There are three parameters to be estimated: persistence ρ, standard

deviation of the innovation σ, and the lower bound of the drift term, $\underline{\mu}$. Since a firm's productivity determines its market value, the first two parameters are disciplined by targeting the estimated autocorrelation of log market value and its coefficient of variation. To pin down ρ, we run the following regression:

$$\ln(ME)_t = A_5 + B_5 \ln(ME)_{t-1} + \varrho_5 \cdot X_{i,t} + \varepsilon_{i,t} \tag{18}$$

where $\ln(ME)$ is the natural logarithm of market equity and the regression controls for year fixed effects. We repeat the same regression using model-simulated firm data, and use the coefficient B_5 in Equation 18 as a target. To pin down σ, we run another regression given by:

$$\ln(ME)_t = A_6 + \varrho_6 \cdot X_{i,t} + \varepsilon_{i,t} \tag{19}$$

where the controls $X_{i,t}$ include leverage, market-to-book ratio, return on assets, cash to assets, tangibility, age, and squared age, as well as year and industry fixed effects. We obtain the residuals from this regression, which are purged from the impact of observables, and calculate its standard deviation and divide it by the sample mean of $\ln(ME)_t$ to obtain a coefficient of variation for log market value. We repeat the same regression using model-generated data and seek to match the estimated coefficient of variation. This leaves us with the final parameter $\underline{\mu}$. An increase in $\underline{\mu}$ increases the growth rate of all firms since it lowers the cost of innovation $X(\mu_{it}, z_{it})$ given in Equation 8 across the board for any productivity drift μ_{it}. Hence, its value is pinned down by targeting the average firm growth rate.

It is worth noting that both ρ and σ are also important determinants of information frictions in our model. It is straightforward to see that information asymmetry and the adverse selection risk it induces are both increasing in σ. The effect of ρ, however, is more nuanced. On the one hand, a high value of ρ implies a more predictable AR(1) process of the target productivity. On the other hand, interestingly, a more persistent AR(1) process also implies that any shocks to the target's productivity would have a larger impact on target valuation, and as a consequence, the adverse selection risk induced by a given level of information asymmetry is actually increasing in ρ. To see the intuition, imagine an acquirer who cannot observe the realization of the target's ongoing innovation and its effect on productivity. This acquirer will be more concerned with the adverse selection risk if the target's productivity is more persistent, because in this case, a negative shock will have a long-lasting effect on the target's productivity and thus hurt the target valuation to a larger extent. As a result, moments related to information asymmetry provide additional identification power for ρ and σ as well.

6. **Exogenous exit rate:** In the model, firms can exit due to endogenous merger decisions, or exogenous shocks. The parameter ψ denotes the latter's probability. In a stationary equilibrium, the entry rate of new firms equals the exit rate. We pin ψ down by targeting the average entry rate of firms in our sample.

7. **Synergy:** γ is the scale parameter of the Cobb-Douglas aggregator that determines the productivity of a merged firm. $\gamma = 2$ would imply that a merger between two identical firms would yield no productivity gains. Values above 2 capture the gains in productivity due to synergy. The value of this parameter is pinned down by targeting the average realized gain in market value (i.e., the combined firm's announcement period abnormal return) and the average merger probability, which are both increasing in γ.

To further demonstrate the identification of model parameters, we report the model-implied sensitivity of moments with respect to parameter values in Online Appendix Table E9 (i.e., the Jacobian matrix). We place the parameters on columns and model-implied moments on rows. To make a clear presentation of the identification, we reorder the parameters and moments in this table so that the parameters are lined up with the main identifying moments around the diagonal (the order of parameters and moments are therefore slightly different from that reported in Table 5). The table reports how model-implied moments change as we perturb the model parameters. We want to emphasize two caveats for this table. First, even though we explain how some moments are more informative in identifying certain parameters, the SMM estimator extracts information from all moments in order to infer parameter values in the estimation process. The mapping between the moments and parameters are clearly not one-to-one. Second, we focus on explaining the economic intuition behind the links between some moments and certain parameters. Even though these moments often carry a large magnitude in the Jacobian matrix, which is consistent with our discussion of identification above, there are occasionally other moments that have an even larger magnitude. This is not a challenge to our discussion of identification, but simply implies that other moments may carry important information regarding the parameter of interest too, which necessitates the use of the SMM estimator.

5.2 Estimation Results

We solve the model numerically and estimate it using SMM. Panel A of Table 5 summarizes the estimated parameter values. The parameter ω, which determines the extent of information asymmetry, is estimated to be 0.697. This estimate implies that there is only 30% of chance that target private information is fully revealed to the acquirer during the merger process. Information asymmetry, therefore, is substantial, and it creates significant barriers in the market for corporate control and hurts efficient asset reallocation in the economy.

Acquirer announcement returns are negatively correlated with equity usage in the data, and the estimated $\Lambda = 0.051$ helps the model fit this data feature fairly well. Though our model does not endogenize the mechanism of this market reaction to the method of payment, a positive Λ captures the costs of using equity payment in takeover offers. It leads to a critical tradeoff in the method of payment: acquirers prefer using equity to hedge adverse selection risk, but equity payment is accepted at a discounted price due to the negative market reaction to the announcement of equity bids.

The parameter of the R&D cost function, $\chi = 58.2$ delivers an average R&D intensity of 5.84%. The persistence of productivity $\rho = 0.863$ is quite high. This delivers a persistence of log firm value of 0.965; very close to the empirical value of 0.944 observed in the data. It suggests that the productivity gains from successful innovation are long-lasting. The standard deviation of innovation $\sigma = 0.374$ is large compared to the mean of $\log(z) = -1.29$. It indicates that there is a non-negligible amount of information asymmetry regarding the innovation of a firm conducted in recent periods.

The estimated probability of exogenous exit ψ is 0.024. Combined with the endogenous exit due to M&A at 0.0165, this delivers a firm entry rate of 4.09%, which is close to the value of 4.57% in our sample, and the overall Compustat entry rate of 5% found in Acemoglu, Akcigit, and Celik (2018).

We estimate γ, the scale factor in merger technology, to be 3.636. A benchmark value of $\gamma = 2$ implies that there is no synergy gain when two identical firms merge together. Our estimate of this parameter indicates that, on average, observed mergers create significant value for the merging pairs.

The parameter α that controls the acquirer's contribution to the combined firm's post-merger productivity is estimated to be 0.644. It implies that the combined firm's productivity is influenced more by the acquirer than by the target. This finding is consistent with the fact that, on average, the merged firm's value depends more on the acquirer's pre-merger value than that of the target.

All parameters are estimated with small standard errors, suggesting that these point estimates are relatively accurate. Two factors contribute to the accuracy. First, most data moments are measured with high precision. Second, the model-implied moments are sensitive to parameter values, a feature that indicates strong identification power of the model.

Panel B of Table 5 reports the model fit by comparing the targeted data moments with their model counterparts.[10] The model manages to match many data features fairly well. For example, in both the model and the data, the equity usage is close to 50% on average, and it increases with the target firm's innovativeness across different deals.

The relation between a firm's innovativeness and its takeover exposure is hump-shaped, with the firms with an intermediate level of innovativeness being acquired most often in the data. When a firm's takeover exposure is regressed on the linear and quadratic terms of its innovativeness, the model produces a positive loading on the linear term and a negative loading on the quadratic term,

[10] To make the coefficients of model-simulated regressions comparable, target's innovation stock in the model is rescaled such that it has the same mean and standard deviation as the target innovation stock in the data as measured by patent citations.

consistent with the data pattern.

The average merger probability is low, and only less than 2% of firms get acquired every year, both in the model and in the data. This value is slightly lower than the average merger rate for the universe of Compustat firms, because firms in our sample are larger on average. Also, the likelihood of deal completion declines with target firm innovativeness, and the model does a good job in capturing this negative association accurately.

The model fit is slightly worse in some other dimensions. For example, the likelihood of being acquired declines slightly faster with firm innovation in the data than that in the model, especially for very large and innovative firms. This discrepancy happens because the model leaves out other factors in the real world that may prevent large, innovative firms from being acquired. For instance, financing costs often make it infeasible for smaller firms to acquire such large and innovative firms in the M&A market. The model also tends to slightly overstate the average realized gain and understate the average value loss in M&A. This is because our model does not feature value destroying mergers due to other frictions (e.g., agency costs). The empirical relation between announcement returns and equity usage allows the model to capture only part of the negative market reaction to merger announcements, and therefore it is not surprising that the model-implied realized gain (value loss) is higher (lower) than that in the data. The model does not fit as tightly in R&D intensity and firm growth as other moments. This happens because there exist some very large, innovative firms in the data, and with Gaussian distributed shocks to productivity, the model simulated distribution cannot perfectly cover these superstar firms. These firms inflate the average R&D intensity and lower the average firm growth rate in the data (i.e., large firms on average have a lower growth rate than smaller firms due to their large base to start with).

But overall, the model is able to match the key features of firm innovation and takeover exposure in the data well, which lends support to the central mechanism we hope to capture in this paper — the information frictions in acquiring innovation.

6 Model Implications

6.1 Reducing Information Frictions

The main goal of our study is to quantify the effect of information frictions on acquisitions of innovation. In our model, conditional on a match, the acquirer cannot observe the true type of the target with probability ω. Intuitively, a lower value of ω can be interpreted as the acquirer being more capable of evaluating the true value of the target's contemporaneous innovation. Conversely, if innovation in the economy is more opaque, it would imply a higher value of ω. If $\omega = 1$, acquirers have no chance to learn the targets' contemporaneous innovation. In this section, we evaluate the effect of information frictions on corporate innovation and firm productivity growth through the channel of an active M&A market.

To do so, we carry out two counterfactual analyses using the estimated model. In the first counterfactual analysis, we fix firms' innovation policy and their cross-sectional distribution as in the baseline model equilibrium. We set $\omega = 0$ and allow the firms to reoptimize their decisions in M&A. Specifically, acquirers choose new optimal offers with perfect information about target innovation, and targets update their acceptance/rejection decisions accordingly. This exercise captures the short-term effect of eliminating information frictions, because we hold firm innovation policy and distribution unchanged. This analysis, therefore, shows the direct effect of information frictions on the M&A market.

In the second counterfactual analysis, we still set $\omega = 0$, but we now allow the firm distribution to evolve endogenously and let firms reoptimize their innovation policy. This exercise captures the long-term effect of eliminating information frictions, and it measures the changes in efficiencies as we move from one equilibrium to another.

To facilitate our analyses, we decompose a firm's market value into two components: its standalone value derived from shutting down its opportunity to participate in the M&A market during its lifetime, and its capitalized expected gains from participating in the M&A market.

Table 6 reports the results. The top four rows present a decomposition of firm market value into its standalone value and the capitalized expected gains from the M&A market, and the bottom five rows show the aggregate-level results. Column one reports the baseline model results. For an average firm in our estimated model, its standalone value accounts for about 94% of the market value, and its capitalized gains from future M&A represents the remaining 6%. An average firm has an R&D intensity around 6% every year, and expands at a rate of 8.6%. On average, 1.65% of firms get acquired each year.

In column two, we evaluate the short-term effect of information frictions as described above. Since we fixed the firms' innovation policy and the cross-sectional distribution as they are in the baseline equilibrium, firm standalone value remains the same. The capitalized expected gains from M&A, however, increase significantly from 0.657 to 1.185, rising from 6% of the market value to more than 10% of firm value. Without information frictions, M&As are more likely to occur in the economy, and the merger rate climbs from 1.65% to 2.30%, representing a 40% increase. Since this counterfactual analysis holds firm innovation policies and the cross-sectional firm distribution unchanged, aggregate implications remain quite similar to those in the baseline.

We then carry out the second counterfactual analysis by allowing firm distribution to evolve and let the firms to reoptimize their innovation policies – in other words, we compute the full long-run general equilibrium effect. Column three confirms that firm innovation activity (i.e., average R&D intensity) and total production (i.e., aggregate output) both increase. Specifically, in the absence of information frictions, firms increase their innovation inputs by about 10%, and the aggregate output rises by 2.65%. Firm growth rate is also boosted by 25 basis points on average.

It is also interesting to note that the long-term effect of information frictions on the M&A market, reported in column three of Table 6, is slightly smaller than the short-term effect reported in column two. For example, the capitalized gains from future M&A is 1.045 in the long run, which is lower than 1.185 in the short run, and the merger probability is 1.95% in the long-run versus 2.3% in the short run. This happens because in the long run, the firm distribution adjusts and many firms of medium innovativeness have been acquired, leaving fewer profitable M&A opportunities in the economy (compared with the short-run effect where the firm distribution is held constant). In general, we can view the long-run effect as a new equilibrium in which information frictions are absent and the short-run effect as part of the trajectory moving from the old equilibrium to a new equilibrium. Eliminating information frictions, therefore, increases firms' capitalized expected gains from the M&A market by 59% in the long run ($59\% = \frac{1.045-0.657}{0.657}$).

Our analyses thus suggest that the inefficiency brought about by information asymmetry is quite sizeable, representing about 3.5% loss in a firm's market value (i.e., $0.035 = \frac{1.045-0.657}{11.048}$). A natural question is how this inefficiency compares with the costs a firm is willing to pay to reduce information asymmetry. For example, is it profitable for an acquiring firm to invest costly efforts in the due diligence process to mitigate the adverse selection risk. To answer this question, we note that the 3.5% value loss arises from information frictions embedded in all M&A deals the firm is expected to be involved in its lifetime. In other words, in order to recover this value loss, the firm needs to invest in due diligence in not only the forthcoming deal but also all future deals. The capitalized costs of such investment can be substantial, close to or even surpassing the 3.5% efficiency loss.

We can also look at the aggregate macroeconomic implications of shutting down information frictions. In the long-run, the absence of information frictions increases aggregate output by 2.65% due to the permanent increase in average firm productivity. The increased efficiency also benefits the representative consumer, and we find that consumption-equivalent welfare increases by 3.02%, implying a substantial social gain.

6.2 Alternative Benchmarks

The benchmark we use in the counterfactual analyses above assume that all information frictions are eliminated. In reality, fully eliminating information asymmetry is probably not achievable. To evaluate how our model implications change when only part of the information frictions can be removed, we experiment with a few alternative benchmarks in this section. We assume that practical policies can only mitigate a fraction (25%, 50%, and 75%) of the information frictions (i.e., set ω to 75%, 50%, and 25% of its estimated value). We present the model implications in these alternative benchmarks in Table 7. As expected, efficiency improvement relative to the alternative benchmarks is smaller. For example, eliminating 50% of information frictions (i.e., moving from baseline to $\omega = 50\% \times \omega^{*}$) increases firms' capitalized expected gains from M&A by 27% in the long-run ($27\% = \frac{0.832}{0.657} - 1$), compared with 59% in the benchmark where all information frictions are eliminated.

Research has shown that, in the course of conducting due-diligence in M&A transactions, lawyers and auditors tend to focus their efforts mainly on the ownership of the company, equity, company debt, as well as the general physical structure of the company. Since most intangible assets are not recognized in financial statements and current accounting rules do not require firms to report separate measures for intangibles, corporate bodies and regulatory authorities inevitably pay little or even no attention to intellectual property (IP) rights. As part of the model's implications on concrete policies, we believe that policies recommending more due diligence on intangible assets and policies that require improved disclosure of intangible assets can help reduce information asymmetry in acquisitions of innovative firms. In fact, due diligence on IP and intangible assets is indeed receiving growing attention among practitioners and corporate managers involved in the M&A transactions in recent years, even though academic research seems to lag behind on this front (see e.g., De Schrijver and Demeyer 2018, Okojie 2018, Negi 2020, and Pohl and Haughey 2021).

6.3 Equity Usage

In the face of the adverse selection risk, acquirers can use equity as payment to bring target shareholders on board and reduce the overpayment risk. Equity usage, therefore, helps mitigate the effect of information frictions and facilitates deal completion. In this section, we carry out a counterfactual analysis in which we force the acquirers to use only cash as payment and examine how it changes our results.

Table 8 presents the findings. We notice that shutting down equity usage reduces firms' capitalized gains from M&A by almost 40% in both the short run and long run (i.e., $\frac{0.404-0.657}{0.657}$), which drops from about 6% of firm value in the baseline model to less than 4% in this counterfactual model with no equity usage. Since acquirers are not allowed to use equity to hedge the adverse selection risk, they become even more conservative in making offers to highly innovative targets, making acquisitions of innovation more challenging.

A less efficient M&A market also has broader impacts on the macroeconomy, and we observe that aggregate output shrinks by 3.1% of its baseline value, which implies a considerable loss in social welfare at 2.40%. Innovation and business dynamism are also adversely affected: Average R&D intensity falls by 11.15% of its base value, and the average firm growth rate falls by 13 basis points.

Overall, this counterfactual analysis sheds light on the value of equity usage in mitigating the negative effect of information frictions in the M&A market.

7 Subsample Estimation and Robustness Checks

7.1 Subsample Estimation

Given the pivotal role of information frictions, we perform a subsample estimation in this section to further explore the cross-sectional variation of our results. We partition firms into two subsamples based on their ex ante measure of information asymmetry and then estimate our model in the two subsamples. The goal is to examine whether our estimated parameters can pick up the different information environments faced by firms in different subsamples.

To construct the ex ante measure of information asymmetry, we obtain Earnings per Share (EPS) forecasts at the firm-analyst level from the Institutional Brokers Estimates System (I/B/E/S) database. We also obtain the realized values of firm EPS from I/B/E/S. For each firm-year observation, following Terry (2017), we focus on the analysts' EPS forecasts in a two-quarter horizon, i.e., from 91 to 180 days before the data release. We first measure each analyst's forecast error for a given firm-year as the difference between each analyst's forecast and the realized value of firm EPS, normalized by the absolute value of realized EPS. Then, for each firm-year observation, we compute the Root Mean Square Error (RMSE) of the forecast errors. We use the RMSE of analyst forecast errors as the ex ante measure of information asymmetry. For each industry in our sample, we calculate the mean of the RMSE of analyst forecast errors. We rank industries according to this mean and split the full sample into high and low RMSE subsamples.

As a first-pass check of these two subsamples, we notice that firms in the high-RMSE subsample use more equity in their M&A transactions compared with those in the low-RMSE subsample (51% vs. 42%). Deal completion rate is also much more sensitive to the level of target innovation in the high-RMSE subsample, compared with the low-RMSE subsample (with a slope of -0.04 vs. -0.03). Firms in the high-RMSE subsample conduct more R&D than those in the low-RMSE subsample (with an R&D intensity of 11% vs. 6%). These patterns are statistically significant and consistent with our expectation of RMSE being an ex ante measure of information asymmetry associated with firm innovation.

We then estimate our model in the two subsamples and report the parameter estimates in Table 9. First and foremost, the estimated ω is 0.75 for the high-RMSE subsample and 0.65 for the low-RMSE subsample. The difference is statistically significant and economically sizeable. Since firms in the high-RMSE subsample invest more on R&D, their estimated costs of innovation χ is also significantly lower than their counterparts in the low-RMSE subsample. Other parameter estimates remain quantitatively close in the two subsamples, implying that the differences in these two subsamples are mainly driven by information asymmetry and the associated innovation.

Lastly, we quantify the effect of information asymmetry in these two subsamples using the estimated model parameters. Table 10 presents the results. As expected, eliminating information asymmetry (i.e., set $\omega = 0$) increases firms' capitalized expected gains from M&A by a much larger margin for the high-RMSE subsample, compared with the low-RMSE subsample ($\frac{1.049-0.633}{0.633} = 66\%$ vs. $\frac{0.953-0.654}{0.654} = 46\%$).

Overall, the subsample estimation results confirm that our estimated parameter ω can indeed pick up the underlying information frictions: for firms with higher ex ante information asymmetry, the acquirer's due diligence reveals only one quarter of the target's private information.

7.2 Extended Model with Target-Side Lack of Information

A key assumption in our baseline model is that target firms are able to observe the realization of their contemporaneous innovation outcome. One possible concern is that even the target firms, sometimes, may not know the realization of their ongoing innovation exactly. To gauge how this assumption affects our results, we extend the baseline model to allow for the possibility that the target has no private information and we explore how it affects firms' M&A and innovation decisions.

In the baseline model, there were two scenarios to consider: the case with imperfect information, where the target knows its own contemporaneous productivity, but the acquirer does not have access to this information; and the case with perfect information, where both the target and the acquirer know the contemporaneous productivity of both firms. In the extended model, we add a third scenario: a case where neither the acquirer nor the target know the contemporaneous productivity of the target firm, and thus both firms have to use rational expectations to formulate their optimal strategies. Let $\Upsilon \in [0,1]$ denote the probability of this new scenario. The model timeline remains similar to that in the baseline, and the Bellman equations are updated to incorporate three scenarios: with probability $(1-\omega)(1-\Upsilon)$, the target's contemporaneous productivity is revealed to the acquirer; with probability $\omega(1-\Upsilon)$, it remains the private information of the target; and with probability Υ, neither the acquirer nor the target know the target's contemporaneous productivity. Details of this model extension are presented in Online Appendix C.

We perform a set of robustness checks with different values of Υ. Specifically, we first use the micro-level data on patent renewals provided by the USPTO to potentially uncover the possible value of Υ, and then we show that the results remain largely robust in a range of Υ that is empirically relevant.

In the United States, inventing firms have to incur large costs to file their patents in the form of application and patent attorney fees on top of the research and development costs to come up with the invention in the first place. Therefore, they only apply for a patent if the expected value of patenting the innovation is above the application costs. Once a patent is granted, firms must also pay a patent maintenance fee in the fourth, eighth, and twelfth years after the grant date to maintain the patent. Unlike the R&D and application costs, patent maintenance fees are quite trivial. However, we observe that some firms choose not to renew their patents despite the trivial maintenance fees. When such an event is observed, we can infer that the firm revised its own valuation of the patent, and decided that its value is below the expectation when the application was filed. The administrative data on patent renewals is thus informative of the frequency of events where a patent turns out to be a "dud" — patents that were initially thought to be worthwhile, but turned out to be of insignificant value ex-post. The frequency of patents that are not renewed in their fourth year is 9.15% for the patents granted to the sample of firms we study. For robustness, we consider low, medium, high, and very high lack of information scenarios, which correspond to $\Upsilon = \frac{9.15\%}{2}$, $\Upsilon = 9.15\%$, $\Upsilon = 2 \times 9.15\%$, and $\Upsilon = 3 \times 9.15\%$, respectively.

Due to space constraints, we report the results predicted by the extended model with different values of Υ in Online Appendix Tables E5 to E8. A higher value of Υ reduces information asymmetry, and we observe that the quantitative importance of information frictions is gradually diminished. However, even if we assume that the target firms do not have private information regarding the realization of their contemporaneous innovation in almost 30% of cases, information frictions still play a crucial role. Specifically, as we set Υ to 27.44% (i.e., three times as large as its calibrated value), eliminating information asymmetry (i.e., set $\omega = 0$) would increase firms' capitalized expected gains from M&A by 36% ($36\% = \frac{1.017}{0.749} - 1$).

These results suggest that the effect of information asymmetry and the resulting adverse selection risk indeed diminish as target firms become less informed of their own innovation outcomes. The inefficiencies induced by information frictions, however, remain quite substantial in a range of Υ that is empirically relevant.

8 Conclusions

Interactions between innovation and the market of M&A have been well studied in the literature. An active M&A market incentivizes some firms to specialize in innovation with the anticipation of being acquired in the future. Acquiring innovation, however, can be challenging because information frictions make it difficult to assess the value and impact of innovation, especially breakthrough innovation. In this paper, we document a robust inverted-U relationship between firm innovativeness and takeover exposure, a positive association between equity usage and target innovativeness, as well as a negative association between deal completion rate and target innovativeness.

We develop and estimate a model of acquiring innovation to quantify the effect of information frictions. We find substantial information frictions between acquirers and targets, and eliminating information frictions can increase firms' expected gains from the M&A market by 59%. A more efficient M&A market stimulates more firm innovation, resulting in higher firm productivity growth and business dynamism, and improves aggregate output and social welfare.

These findings highlight the importance of the M&A market as a channel of optimal resource allocation and technological progress. Policies that can alleviate the friction in this market would not only increase the number of mergers, but also boost innovation across firm boundaries as firms innovate more in anticipation of future merger opportunities. At the same time, larger firms which are relatively inefficient in R&D can reduce their in-house R&D expenditure, which makes each dollar spent on innovation more effective. Business dynamism also improves, thanks to the higher turnover of firms, and all firms, regardless of size and innovativeness, experience an increase in firm value. Given these significant positive effects, we believe future analyses targeted towards uncovering new ways to reduce information frictions to be socially worthwhile.

References

ACEMOGLU, D., U. AKCIGIT, AND M. A. CELIK (2018): "Young, restless and creative: Openness to disruption and creative innovations," *NBER Working Paper*.

AKERLOF, G. A. (1970): "The market for "lemons": Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, 84(3), 488–500.

ALBUQUERQUE, R., AND E. SCHROTH (2010): "Quantifying private benefits of control from a structural model of block trades," *Journal of Financial Economics*, 96(1), 33–55.

——— (2015): "The value of control and the costs of illiquidity," *The Journal of Finance*, 70, 1405–1455.

ARCAND, J. L., E. BERKES, AND U. PANIZZA (2015): "Too much finance?," *Journal of Economic Growth*, 20(2), 105–148.

AUSUBEL, L. M., P. CRAMTON, AND R. DENECKERE (2002): "Bargaining with incomplete information," *Chapter 50 of Handbook of Game Theory (R. Aumann and S. Hart, eds.), Amsterdam: Elsevier Science B.V.*, 3.

BATT, R. J., AND C. TERWIESCH (2017): "Early task initiation and other load-Adaptive mechanisms in the emergency department," *Management Science*, 63(11), 3531–3551.

BAZZI, S., A. GADUH, A. D. ROTHENBERG, AND M. WONG (2016): "Skill transferability, migration, and development: Evidence from population resettlement in indonesia," *American Economic Review*, 106, 2658–2698.

BENA, J., AND K. LI (2014): "Corporate innovations and mergers and acquisitions," *The Journal of Finance*, 69, 1923–1960.

BHAGAT, S., M. DONG, D. HIRSHLEIFER, AND R. NOAH (2005): "Do tender offers create value? New methods and evidence," *Journal of Financial Economics*, 76, 3–60.

CAVENAILE, L., M. A. CELIK, AND X. TIAN (2021): "The dynamic effects of antitrust policy on growth and welfare," *Journal of Monetary Economics*.

CUNNINGHAM, C., F. EDERER, AND S. MA (2019): "Killer acquisitions," *Working paper*.

DAVID, J. (2021): "The aggregate implications of mergers and acquisitions," *The Review of Economic Studies*, 88(4), 1796–1830.

DE SCHRIJVER, S., AND M. DEMEYER (2018): "IP due diligence in M&A transactions," *Corporate Live Wire (https://www.corporatelivewire.com/top-story.html?id=ip-due-diligence-in-ma-transactions)*.

DIMOPOULOS, T., AND S. SACCHETTO (2017): "Merger activity in industry equilibrium," *Journal of Financial Economics*, 126(1), 200–226.

ECKBO, B. E., R. M. GIAMMARINO, AND R. L. HEINKEL (1990): "Asymmetric information and the medium of exchange in takeovers: Theory and tests," *The Review of Financial Studies*, 3(4), 651–675.

EISFELDT, A. L., AND A. RAMPINI (2008): "Managerial incentives, capital reallocation, and the business cycle," *Journal of Financial Economics*, 87, 177–199.

ERICKSON, T., AND T. WHITED (2002): "Two-step GMM estimation of the errors-invariables model using high-order moments," *Econometric Theory*, 18, 776–799.

FISHMAN, M. J. (1989): "Preemptive bidding and the role of the medium of exchange in acquisitions," *The Journal of Finance*, 44(1), 41–57.

GORBENKO, A. S., AND A. MALENKO (2014): "Strategic and financial bidders in takeover auctions," *The Journal of Finance*, 69, 2513–2555.

HALL, B. H. (1987): "The relationship between firm size and firm growth in the U.S. manufacturing sector," *Journal of Industrial Economics*, 35, 583–600.

HALL, B. H., A. B. JAFFE, AND M. TRAJTENBERG (2001): "The NBER patent citation data file: Lessons, insights and methodological tools," *NBER Working Paper*.

HANSEN, R. G. (1987): "A theory for the choice of exchange medium in the market for corporate control," *Journal of Business*, 60(1), 75–95.

HIGGINS, M. J., AND D. RODRIGUEZ (2006): "The outsourcing of R&D through acquisitions in the pharmaceutical industry," *Journal of Financial Economics*, 80, 351–383.

JOVANOVIC, B., AND P. L. ROUSSEAU (2002): "The Q-theory of mergers," *American Economic Review*, 92(2), 198–204.

KESAVAN, S., B. R. STAATS, AND W. GILLAND (2014): "Volume flexibility in services: The costs and benefits of flexible labor resources," *Management Science*, 60(8), 1884–1906.

LEE, B.-S., AND B. F. INGRAM (2010): "Simulation estimation of time series models," *Journal of Econometrics*, 47(1), 197–205.

LEVINE, O. (2017): "Acquiring growth," *Journal of Financial Economics*, 126(2), 300–319.

LI, D., L. A. TAYLOR, AND W. WANG (2018): "Inefficiencies and externalities from opportunistic acquirers," *Journal of Financial Economics*, 130(2), 265–290.

LIND, J. T., AND H. MEHLUM (2010): "With or without U? The appropriate test for a U-shaped relationship: Practitioners' corner," *Oxford Bulletin of Economics and Statistics*, 72(1), 109–118.

NEGI, S. (2020): "IP becomes even more integral to M&A due diligence," *Copperpod Intellectural Property (https://www.copperpodip.com/post/ip-becomes-even-more-integral-to-m-a-due-diligence)*.

OKOJIE, Y. (2018): "The importance of IP due diligence in mergers and acquisitions," *SPA Ajibade Co.*

PHILLIPS, G. M., AND A. ZHDANOV (2013): "R&D and the incentives from merger and acquisition activity," *Review of Financial Studies*, 26, 34–78.

POHL, S., AND P. HAUGHEY (2021): "One size does not fit all: Tailoring IP due diligence to the transaction," *Association of Corporate Counsel (https://www.acc.com/resource-library/one-size-does-not-fit-all-tailoring-ip-due-diligence-transaction)*.

RHODES-KROPF, M., AND D. T. ROBINSON (2008): "The market for mergers and the boundaries of the firm," *The Journal of Finance*, 63(3), 1169–1211.

RHODES-KROPF, M., D. T. ROBINSON, AND S. VISWANATHAN (2005): "Valuation waves and merger activity: The empirical evidence," *Journal of Financial Economics*, 77(3), 561–603.

RHODES-KROPF, M., AND S. VISWANATHAN (2004): "Market valuation and merger waves," *The Journal of Finance*, 59(6), 2685–2718.

RODRIK, D. (2016): "Premature deindustrialization," *Journal of Economic Growth*, 21(1), 1–33.

SEVILIR, M., AND X. TIAN (2011): "Acquiring innovation," *Unpublished working paper. Indiana University*.

SHLEIFER, A., AND R. W. VISHNY (2003): "Stock market driven acquisitions," *Journal of Financial Economics*, 70(3), 295–311.

STREBULAEV, I. A., AND T. M. WHITED (2012): "Dynamic models and structural estimation in corporate finance," *Foundations and Trends in Finance*, 6, 1–163.

TAN, T. F., AND S. NETESSINE (2014): "When does the devil make work? An empirical study of the impact of workload on worker productivity," *Management Science*, 60(6), 1574–1593.

TERRY, S. (2017): "The macro impact of short-termism," *Working paper*.

VLADIMIROV, V. (2015): "Financing bidders in takeover contests," *Journal of Financial Economics*, 117(3), 534–557.

WANG, X. (2018): "Catering innovation: Entrepreneurship and the acquisition market," *Working Paper*, (18-27).

WARUSAWITHARANA, M. (2008): "Corporate asset purchases and sales: Theory and evidence," *Journal of Financial Economics*, 87, 471–497.

YANG, L. (2008): "The real determinants of asset sales," *The Journal of Finance*, 63(5), 2231–2262.

Figures and Tables

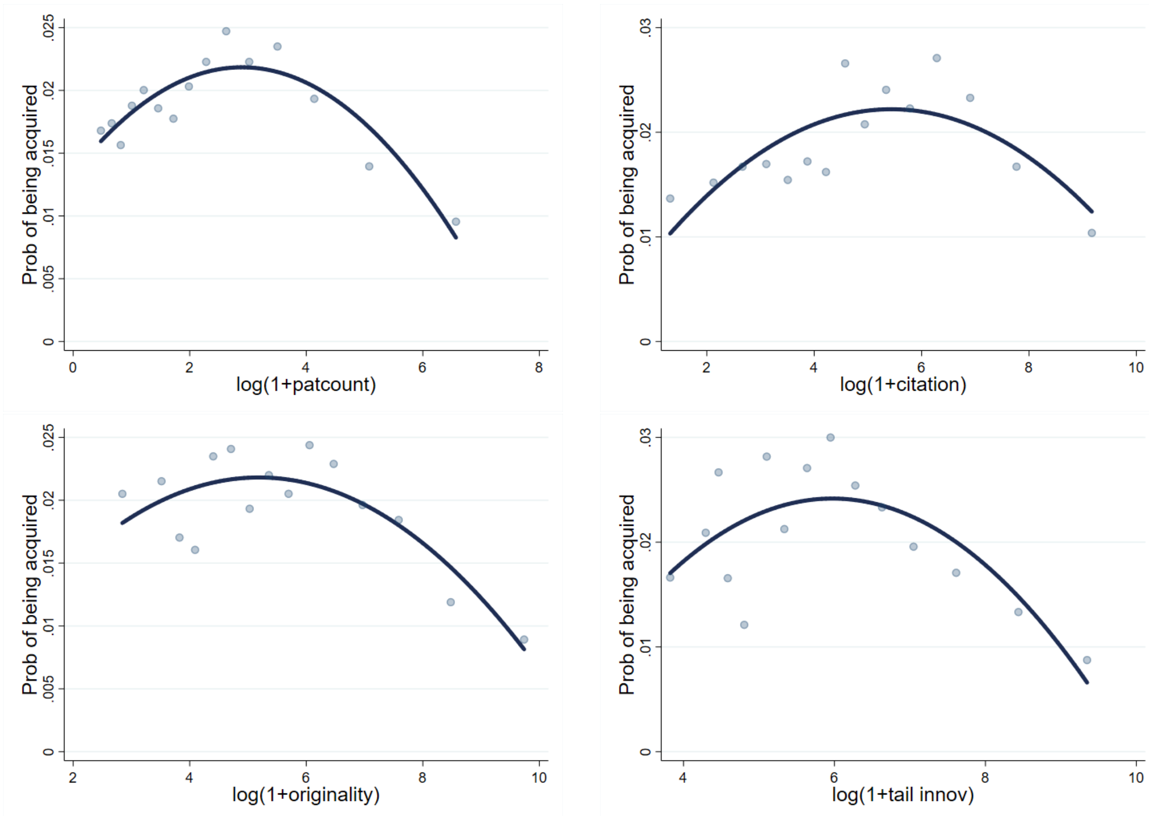

FIGURE 1: FIRM INNOVATION AND THE PROBABILITY OF BEING ACQUIRED

Notes: This figure illustrates the relation between a firm's past innovation stock and its probability of being acquired next year without any controls. The measures of firm innovation are patent count, patent citations, originality, and tail innovations (top 10%). Variables are defined in Table 1. Dots represent the average probability of being acquired for firms in intervals partitioned by their innovation, and the curves represent the fitted value generated by quadratic regressions.

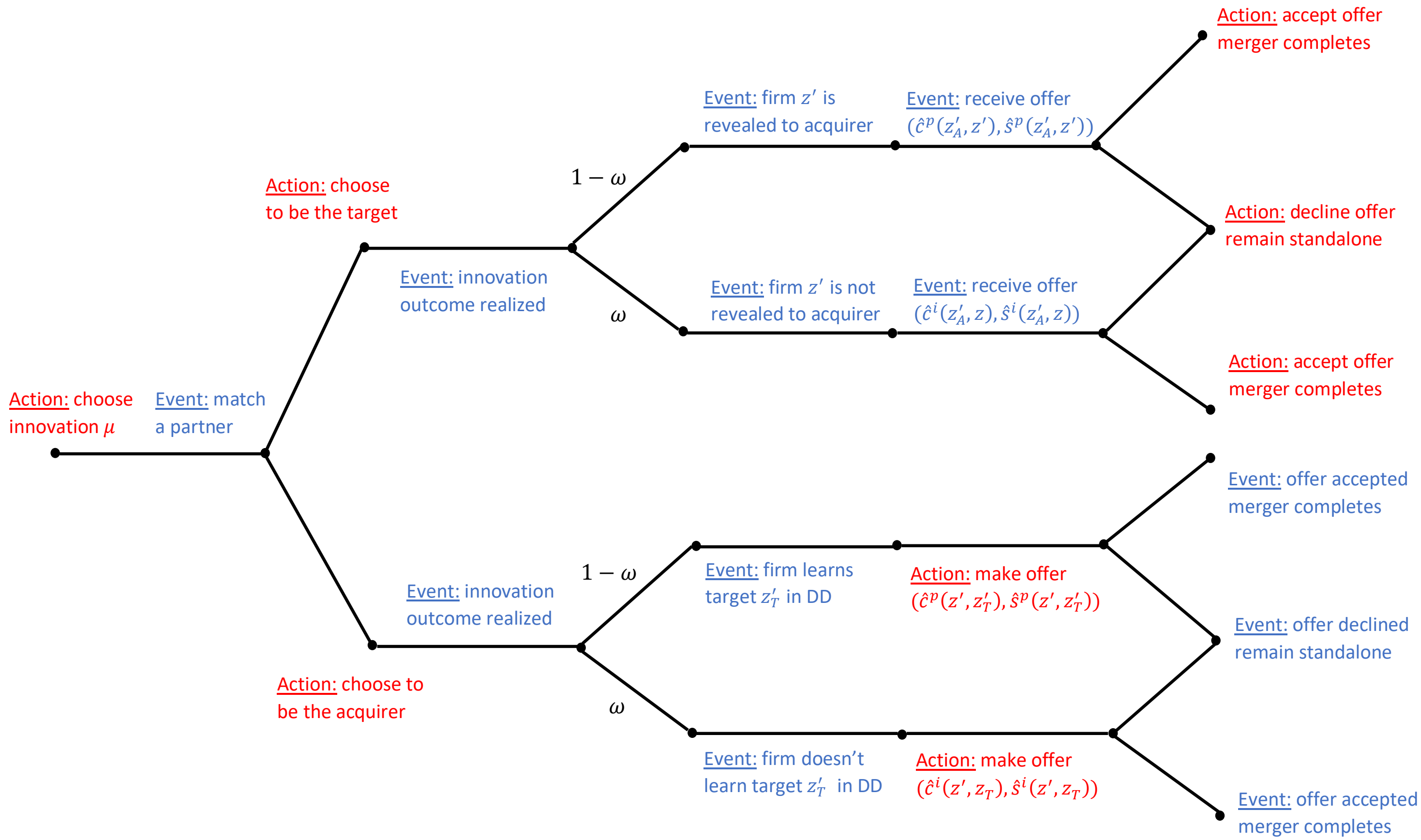

FIGURE 2: MODEL TIMELINE

Notes: This figure illustrates the model timeline within a given period t. It describes the events (in blue) and actions (in red) involving a firm in our model as well as its information set. At the beginning of period t, the firm chooses its innovation level, and then randomly matches with a potential merger partner. The firm chooses its potential role in the merger, which generates two branches: (1) If the firm acts as the target (the upper branch), it learns the contemporaneous productivity of itself and the acquirer (i.e., $z^{'}$ and $z^{'}_A$) immediately after the innovation outcomes are realized. The acquirer then conducts due diligence (DD). With probability $1-\omega$, the acquirer learns the firm's contemporaneous productivity, $z^{'}$, and with probability ω, the acquirer learns nothing new. The firm then receives an offer from the acquirer and decides to accept or decline it. (2) If the firm acts as the acquirer (the lower branch), it learns its own contemporaneous productivity immediately after the innovation outcomes are realized. It, however, does not observe the target's new productivity. The firm then conducts due diligence which reveals the target's type with a probability of $1-\omega$. Based on the firm's information set after due diligence, it makes an offer to the target, which may be accepted or declined. If an offer is accepted, then two matched firms merge together and the target firm exits; if an offer is declined, then both firms remain standalone.

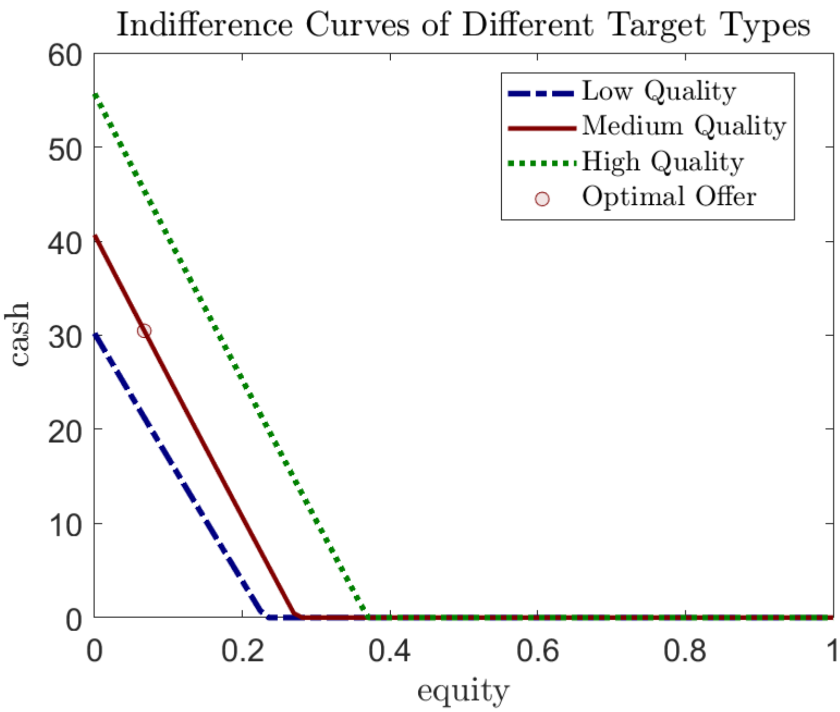

FIGURE 3: TARGET INDIFFERENCE CURVES

Notes: This figure illustrates an example where an acquirer with type z_A meets a target with pre-innovation productivity $\tilde{z}_T$. The true productivity of the target, z_T, is not public information, so the acquirer must form rational expectations based on the innovation policy of the target and the law of motion for productivity. For simplicity, we consider a discretized version with three possible types for the target: $z_T^h > z_T^m > z_T^l$ with $z_T^m = \tilde{z}_T$. The acquirer knows how each of the three types would react to any offer (c, s). The figure depicts the indifference curves of each target type along which they are indifferent between accepting or rejecting the merger offer. Any offer above the curve would be accepted. In this example, an offer that is acceptable to z_T^h would be acceptable to both z_T^m and z_T^l, but the opposite is not true. Given this information, the acquirer chooses the amount of cash and equity in the offer, (c, s), to maximize its expected value.

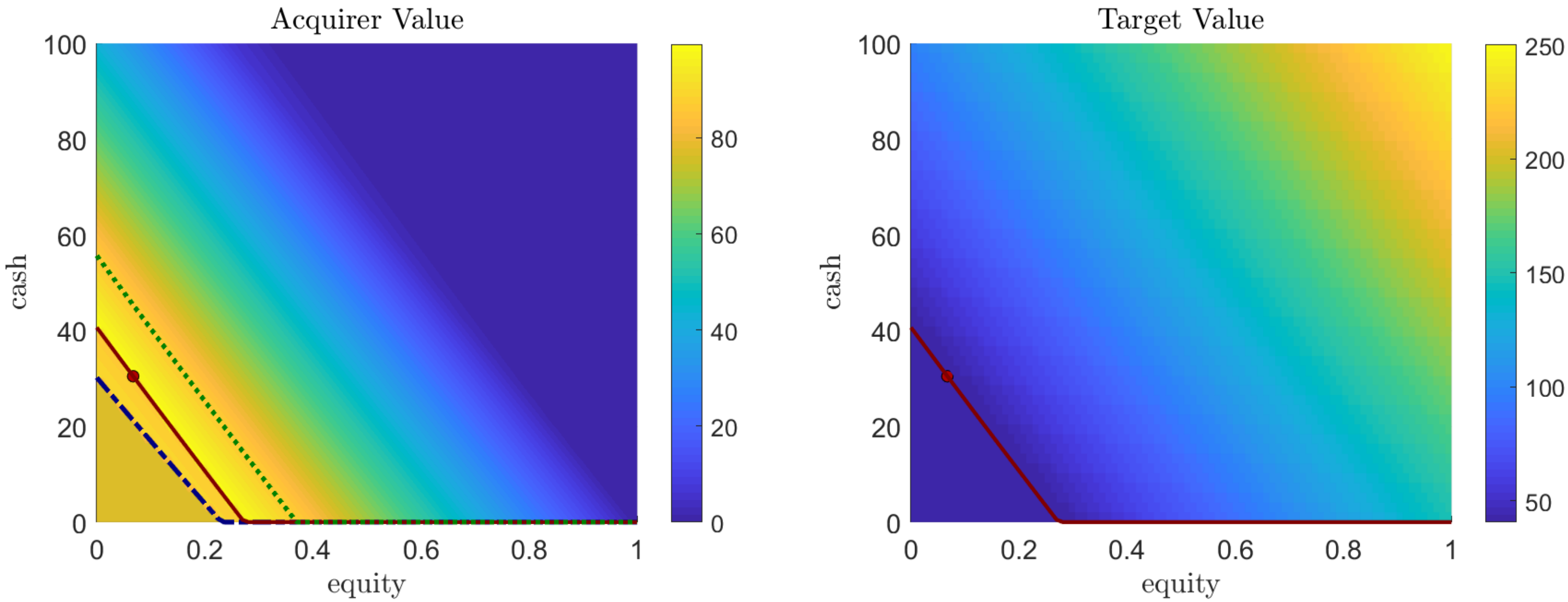

FIGURE 4: ACQUIRER AND TARGET VALUE

Notes: The left panel of this figure displays the expected value of the acquirer as a heat map for the example in Figure 3. Given that there are only three types of targets, the optimal offer must lie on one of the target indifference curves, because any deviation from the indifference curve implies overpayment (the acquirer can pay less but still get the same set of targets to accept the offer). In this example, the value function of the acquirer is maximized at the red dot, and the offer contains both cash and equity. This offer is accepted by the low and medium target types, but rejected by the high type. The right panel of this figure displays the value function of the medium type target as a heat map. The value function is increasing in c and s above the indifference curve. However, it is flat below the indifference curve where its level is equal to the target's continuation value (reservation value). Therefore, the target type that marginally accepts the offer is only paid its outside option, whereas the low type is paid in excess of its outside option, and collects information rents. This simple logic carries over to a discretization with more types, as well as a continuous z_T.

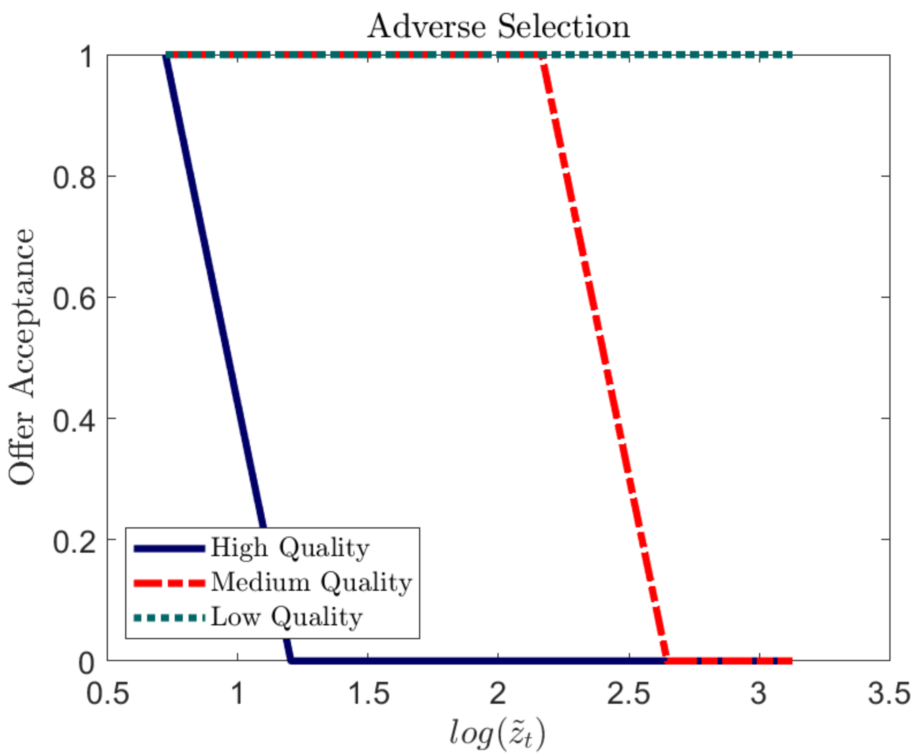

FIGURE 5: OFFER ACCEPTANCE: ADVERSE SELECTION

Notes: This figure depicts the deal acceptance policy functions of the three post-innovation types defined in Figure 3 for multiple values of observed target innovativeness, $\tilde{z}_T$. Acceptance is denoted as 1. In this example, it is seen that when $\tilde{z}_T$ is low, the acquirer makes an offer that would be accepted by all types. This indicates that the expected value of the merger to the acquirer is so high that it does not mind overpaying the medium and low types. As $\tilde{z}_T$ increases, the uncertainty becomes less tolerable. The adverse selection problem is aggravated, and the acquirer makes a more timid offer, which is rejected by the high type, but accepted by the medium and low types. For very high values of $\tilde{z}_T$, only firms with the lowest type accept the acquirer's offer – a market for lemons.

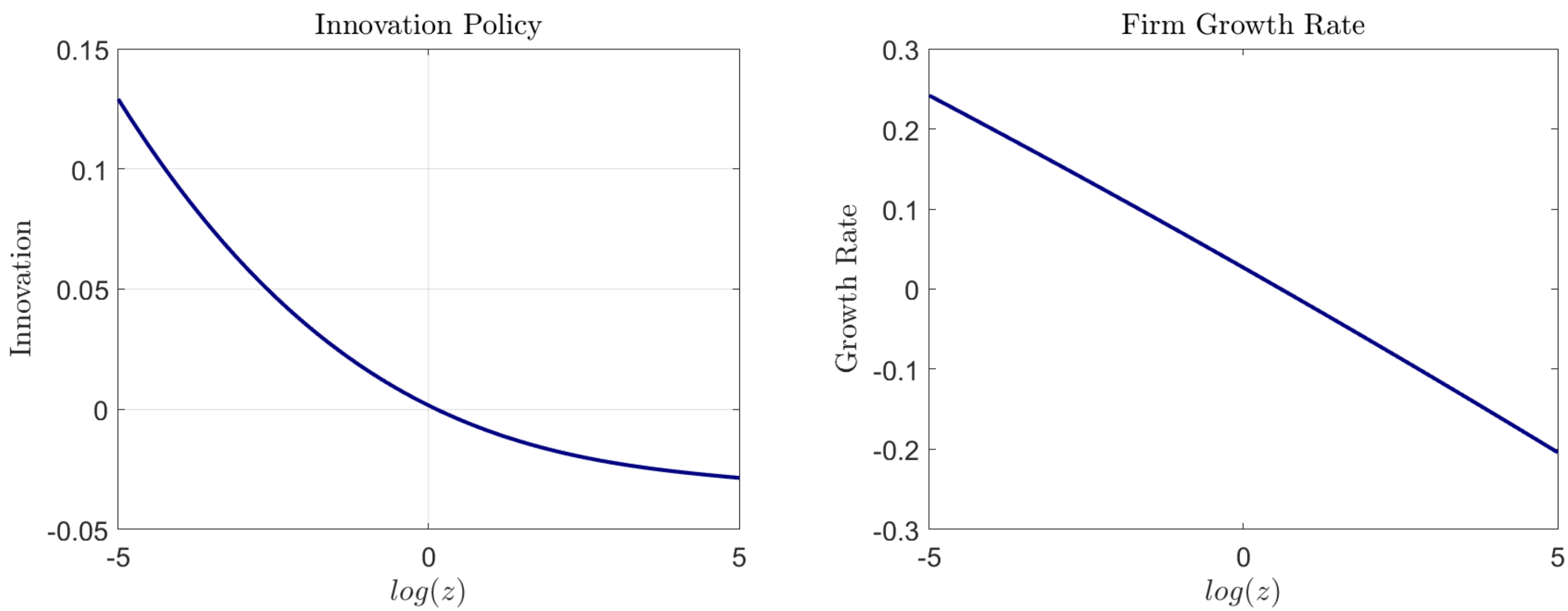

FIGURE 6: INNOVATION POLICY AND FIRM GROWTH

Notes: This figure depicts the optimal innovation policy function, and the implied expected firm growth rate for firms with different productivity levels z in the left and right panels, respectively. Consistent with the empirical findings in previous literature that reject Gibrat's law, smaller firms innovate more intensively and grow faster. The existence of an active M&A market creates an extra incentive for small firms to innovate more and grow faster, as becoming more productive makes them more valuable to potential acquirers. Large firms can choose a lower innovation intensity, as they have the option of improving their productivity by acquiring productive small firms rather than spending too much on in-house R&D.

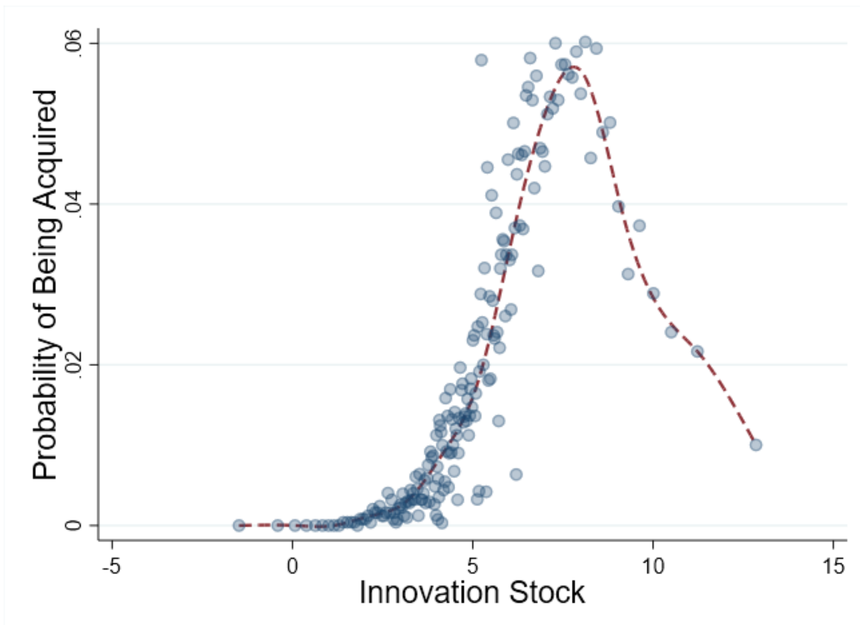

FIGURE 7: PROBABILITY OF BEING ACQUIRED AND FIRM INNOVATIVENESS

Notes: This figure depicts the relationship between firm innovativeness (measured as the innovation stock of the firm) and the probability of being acquired. Scattered dots represent the simulated data from the estimated model, grouped into bins of innovation stock. The red dashed line is the fitted curve. When firm innovativeness is quite low, the benefit from M&A is too low for an acquirer to make an offer, and thus the probability of being acquired is zero. As firm innovativeness increases, so does its potential contribution as a target in M&A. Thus, the firm is more likely to receive offers, and the probability of being acquired increases. The probability of being acquired, however, starts decreasing as the firm becomes highly innovative due to an intensified adverse selection risk.

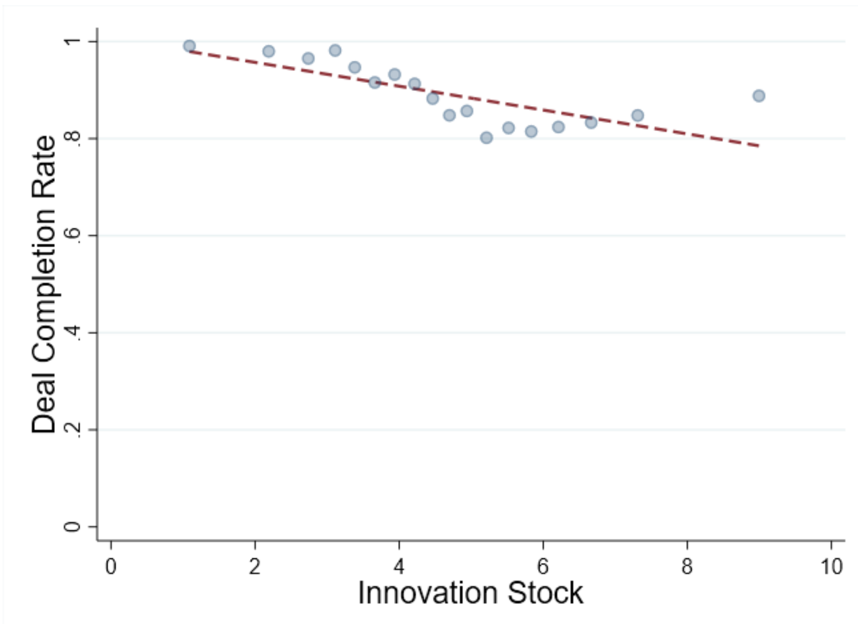

FIGURE 8: DEAL COMPLETION AND TARGET INNOVATIVENESS

Notes: This figure displays a negative relation between the likelihood of deal completion and the target innovativeness (measured as the innovation stock of the target). Scattered dots represent the simulated data from the estimated model, grouped into bins of innovation stock. The red dashed line is the fitted curve. Deal completion rate drops as the target firm innovativeness increases, because acquirers are more reluctant to offer a high price in face of adverse selection risk that aggravates with target firm innovativeness.

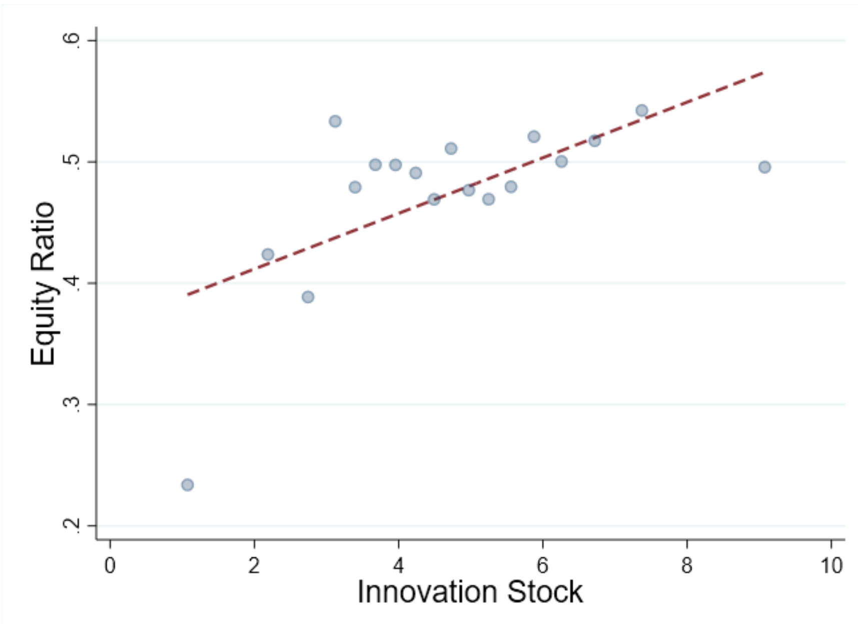

FIGURE 9: OFFER COMPOSITION AND TARGET INNOVATIVENESS

Notes: This figure shows a positive relationship between the equity ratio in an offer and the target innovativeness (measured as the innovation stock of the target). Scattered dots represent the simulated data from the estimated model, grouped into bins of innovation stock. The red line is the fitted curve. Since the acquirers cannot always perfectly observe the target's true productivity under information asymmetry, they hedge the adverse selection risk by using equity in their offers. Adverse selection risk increases with target firm innovation, and this leads to an increased usage of equity as the targets become more innovative.

TABLE 1: VARIABLE DEFINITIONS

Notation	Variable	Definition
$Patent\ count$	innovation quantity (patent count)	$ln\,(1+\text{patcount})$
$Citations$	innovation quality (patent citations)	$ln\,(1+\text{citation})$
$Originality$	innovation originality	$ln\,(1+\text{originality})$
$Breakthrough$	breakthrough innovation	$ln\,(1+\text{tail innov})$
$Size$	logarithm of market equity	$ln\,(ME)$
$Leverage$	market leverage	$\frac{dltt+dlc}{dltt+dlc+ME+PSLV-txditc}$
MB	market-to-book equity ratio	$\frac{ME}{BE}$
ROA	return on assets	$\frac{ni}{at}$
$Cash$	cash holdings	$\frac{che}{at}$
$R\&D$	research and development	$\frac{xrd}{at}$
$Tangibility$	asset tangibility	$\frac{ppeng}{at}$
Div	dividend dummy	$I_{div>0}$
Age	firm age	age
RRV	misvaluation measure	as in Rhodes-Kropf, Robinson, and Viswanathan (2005)
$Diversification$	diversifying dummy	$I_{sic_{acq}\neq sic_{tar}}$
$RelSize_{tar}$	target size relative to the acquirer	$\frac{ME_{tar}}{ME_{acq}}$
$PrcOfStk$	fraction of equity in bids	$\frac{Equity\ payment}{Total\ payment}$
$ln(Assets)$	logarithm of assets	$ln\,(at)$
$CombRet$	the combined firm abnormal return	combined firm 3-day CAR plus 30-day run-up

Notes: This table provides variable definitions. The sample consists of all U.S. listed firms, covering the period of 1980-2006. All innovation variables are constructed as the stock of past innovations following the perpetual inventory method with a depreciation rate of 6%. Patent count is the number of patents applied for by a firm in a given year. Patent citations are total citations received by patents applied for by a firm in a given year, which are corrected for truncation and technology class biases following Hall, Jaffe, and Trajtenberg (2001). Originality is the dispersion of technology classes cited by the firm's patents as described in Hall, Jaffe, and Trajtenberg (2001). Breakthrough innovations are defined as the number of patents that are in the top 10% of all patents according to the number of citations received among all patents applied for in that year as in Acemoglu, Akcigit, and Celik (2018).

TABLE 2: FIRM INNOVATION AND THE PROBABILITY OF BEING ACQUIRED

Dependent var:	Probability of being acquired (annual)			
Innovation var:	Patent count	Citations	Originality	Breakthrough
Innov	0.0036**	0.0060***	0.0061***	0.0179***
	(2.39)	(3.57)	(2.91)	(4.50)
$Innov^2$	-0.0009***	-0.0007***	-0.0007***	-0.0016***
	(-4.46)	(-4.12)	(-4.18)	(-5.03)
$ln(Assets)$	0.0025***	0.0020***	0.0026***	0.0018***
	(4.03)	(4.09)	(3.79)	(2.69)
Leverage	-0.0065***	-0.0060***	-0.0054*	-0.005
	(-3.01)	(-2.80)	(-1.93)	(-1.46)
MB	-0.0001	-0.0001	-0.0001	-0.0001
	(-0.51)	(-0.50)	(-0.86)	(-0.86)
ROA	0.0015**	0.0018**	0.0019***	0.0033***
	(2.47)	(2.50)	(3.49)	(2.89)
Cash	-0.0001	0.0001	-0.0003	0.0043
	(-0.02)	(0.01)	(-0.05)	(0.54)
R&D	0.0182***	0.0174***	0.0181***	0.0183**
	(5.23)	(4.28)	(5.32)	(2.37)
Tangibility	-0.006	-0.0053	-0.0072	-0.0061
	(-1.56)	(-1.34)	(-1.68)	(-1.07)
Div	-0.0080***	-0.0078***	-0.0085***	-0.0085***
	(-3.73)	(-3.68)	(-3.42)	(-3.04)
Age	0.0010***	0.0010***	0.0010***	0.0012***
	(3.92)	(3.93)	(3.98)	(3.79)
Age^2	-0.0000***	-0.0000***	-0.0000***	-0.0000***
	(-3.35)	(-3.47)	(-3.29)	(-3.18)
Const	0.0085**	0.0012	-0.0001	-0.0419***
	(2.51)	(0.20)	(-0.01)	(-3.23)
Industry F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	59822	57885	49348	35433
adjusted R^2	0.006	0.005	0.006	0.007

Notes: This table reports the regression results from Equation 1. *Innov* is the firm's innovation, measured by its patent count (column 1), patent citations (column 2), originality (column 3), and breakthrough innovations (column 4). All innovation variables are constructed as the stock of past innovations following the perpetual inventory method with a depreciation rate of 6%. Patent count is the number of patents applied for by a firm in a given year. Patent citations are total citations received by patents applied for by a firm in a given year, which are corrected for truncation and technology class biases following Hall, Jaffe, and Trajtenberg (2001). Originality is the dispersion of technology classes cited by the firm's patents as described in Hall, Jaffe, and Trajtenberg (2001). Breakthrough innovations are defined as the number of patents that are in the top 10% of all patents according to the number of citations received among all patents applied for in that year as in Acemoglu, Akcigit, and Celik (2018). Other variable definitions are in Table 1. Standard errors are clustered by the firm's industry (Fama-French 48-industry classification).

TABLE 3: FIRM INNOVATION AND THE METHOD OF PAYMENT

Dependent var:	Equity share in acquisition offer			
Innovation var:	Patent count	Citations	Originality	Breakthrough
$Innov_{tar}$	0.0237	0.0243**	0.0265*	0.0425*
	(1.64)	(2.18)	(1.80)	(1.94)
RRV_{acq}	0.1060***	0.1070***	0.0916***	0.1138***
	(3.36)	(3.65)	(3.02)	(3.88)
$Innov_{acq}$	0.0021	0.0019	-0.0077	0.0074
	(0.22)	(0.30)	(-1.16)	(0.83)
$Size_{acq}$	-0.0646***	-0.0684***	-0.0489***	-0.0776***
	(-4.74)	(-5.34)	(-4.89)	(-3.36)
MB_{acq}	0.0182	0.0212*	0.0183	0.0002
	(1.62)	(1.75)	(1.53)	(0.01)
$Leverage_{acq}$	-0.0082	0.0025	-0.0468	-0.1368
	(-0.08)	(0.02)	(-0.42)	(-1.22)
ROA_{acq}	-0.3549***	-0.3668***	-0.3428***	-0.3272
	(-2.84)	(-2.91)	(-3.06)	(-1.45)
RRV_{tar}	0.0923***	0.0961***	0.0898***	0.1197***
	(3.87)	(4.05)	(4.13)	(3.96)
$RelSize_{tar}$	0.0031	0.0029	0.0094	-0.0003
	(0.27)	(0.26)	(1.23)	(-0.01)
MB_{tar}	0.0079	0.0028	0.0083	-0.0026
	(0.71)	(0.22)	(0.82)	(-0.20)
$Leverage_{tar}$	0.1192	0.1169	0.1027	0.0710
	(1.30)	(1.24)	(1.04)	(0.77)
ROA_{tar}	0.0375	0.0302	0.0047	0.0901
	(0.44)	(0.39)	(0.06)	(1.16)
$Diversification$	-0.0255	-0.0190	0.0158	-0.0226
	(-0.51)	(-0.38)	(0.26)	(-0.40)
$Const$	1.4836***	1.4976***	1.2204***	1.6832***
	(5.30)	(5.66)	(5.53)	(3.33)
Industry F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	605	590	530	410
adjusted R^2	0.253	0.255	0.245	0.279

Notes: This table reports the regression results from Equation 2. *Innov* is the firm's innovation, measured by its patent count (column 1), patent citations (column 2), originality (column 3), and breakthrough innovations (column 4). All innovation variables are constructed as the stock of past innovations following the perpetual inventory method with a depreciation rate of 6%. Patent count is the number of patents applied for by a firm in a given year. Patent citations are total citations received by patents applied for by a firm in a given year, which are corrected for truncation and technology class biases following Hall, Jaffe, and Trajtenberg (2001). Originality is the dispersion of technology classes cited by the firm's patents as described in Hall, Jaffe, and Trajtenberg (2001). Breakthrough innovations are defined as the number of patents that are in the top 10% of all patents according to the number of citations received among all patents applied for in that year as in Acemoglu, Akcigit, and Celik (2018). Other variable definitions are in Table 1. Standard errors are clustered by the acquirer's industry (Fama-French 48-industry classification).

TABLE 4: FIRM INNOVATION AND THE PROBABILITY OF DEAL COMPLETION

Dependent var:	Offer acceptance indicator			
Innovation var:	Patent count	Citations	Originality	Breakthrough
$Innov_{tar}$	-0.0362***	-0.0295***	-0.0287***	-0.0441***
	(-3.66)	(-3.64)	(-3.04)	(-3.32)
RRV_{acq}	-0.0203	-0.0190	-0.0064	-0.0305
	(-0.84)	(-0.78)	(-0.25)	(-0.98)
$Innov_{acq}$	0.0030	0.0029	0.0041	-0.0054
	(0.36)	(0.49)	(0.71)	(-0.83)
$Size_{acq}$	0.0294**	0.0304**	0.0293**	0.0332**
	(2.30)	(2.52)	(2.42)	(2.16)
MB_{acq}	0.0035	0.0042	0.0089	-0.0001
	(0.38)	(0.45)	(0.91)	(-0.00)
$Leverage_{acq}$	-0.0999	-0.0992	-0.0737	-0.1720
	(-1.31)	(-1.28)	(-0.92)	(-1.60)
ROA_{acq}	-0.1498	-0.1499	-0.0975	-0.1238
	(-1.35)	(-1.35)	(-0.86)	(-0.78)
RRV_{tar}	0.0398**	0.0412**	0.0249	0.0719***
	(2.07)	(2.10)	(1.21)	(3.15)
$RelSize_{tar}$	-0.0559***	-0.0570***	-0.0456***	-0.0789**
	(-4.11)	(-4.19)	(-3.09)	(-2.44)
MB_{tar}	-0.0190**	-0.0199**	-0.0134	-0.0234**
	(-2.23)	(-2.26)	(-1.43)	(-2.28)
$Leverage_{tar}$	-0.0097	-0.0161	-0.0525	-0.0134
	(-0.16)	(-0.26)	(-0.81)	(-0.18)
ROA_{tar}	-0.1073	-0.1235*	-0.1583*	-0.0727
	(-1.46)	(-1.66)	(-1.95)	(-0.83)
$Diversification$	-0.0223	-0.0251	-0.0246	0.0036
	(-0.75)	(-0.83)	(-0.77)	(0.10)
$Const$	0.5974*	0.6178*	0.5081	0.7062*
	(1.75)	(1.86)	(1.24)	(1.86)
Industry F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	678	663	577	457
adjusted R^2	0.136	0.134	0.135	0.146

Notes: This table reports the regression results from Equation 3. *Innov* is the firm's innovation, measured by its patent count (column 1), patent citations (column 2), originality (column 3), and breakthrough innovations (column 4). All innovation variables are constructed as the stock of past innovations following the perpetual inventory method with a depreciation rate of 6%. Patent count is the number of patents applied for by a firm in a given year. Patent citations are total citations received by patents applied for by a firm in a given year, which are corrected for truncation and technology class biases following Hall, Jaffe, and Trajtenberg (2001). Originality is the dispersion of technology classes cited by the firm's patents as described in Hall, Jaffe, and Trajtenberg (2001). Breakthrough innovations are defined as the number of patents that are in the top 10% of all patents according to the number of citations received among all patents applied for in that year as in Acemoglu, Akcigit, and Celik (2018). Other variable definitions are in Table 1. Standard errors are clustered by the acquirer's industry (Fama-French 48-industry classification).

TABLE 5: BASELINE MODEL PARAMETERS AND TARGET MOMENTS

A. Parameter estimates

Parameter	Description	Value	Std. Dev.
ω	information asymmetry	0.697	0.010
Λ	market reaction to equity usage	0.051	0.001
χ	innovation cost	58.194	0.727
ρ	productivity persistence	0.863	0.003
σ	innovation std. dev.	0.374	0.002
$\underline{\mu}$	lower bound of productivity drift	-0.032	0.001
ψ	exogenous exit rate	0.024	0.001
γ	merger technology, scale	3.636	0.042
α	merger technology, acquirer share	0.644	0.007

B. Moments

Target Moments	Model	Data	Std. Err.
average equity ratio	49.63%	47.17%	1.87%
loading of equity ratio on innov stock ($B_2 \times 100$ in Eq 2)	2.282	2.432	0.282
loading of merger prob. on innov stock ($B_1 \times 100$ in Eq 1)	0.865	0.602	0.011
loading of merger prob. on innov stock^2 ($C_1 \times 100$ in Eq 1)	-0.020	-0.068	0.002
average merger probability	1.65%	1.91%	0.06%
loading of deal completion on innov stock ($B_3 \times 100$ in Eq 3)	-2.358	-2.953	0.233
average realized gain	5.32%	3.16%	0.54%
average value loss	0.68%	1.09%	0.04%
average R&D intensity	5.84%	8.50%	1.77%
autocorr. of $ln(ME)$ (B_5 in Eq 18)	0.965	0.944	0.001
coefficient of variation of $ln(ME)$ from Eq 19	0.210	0.364	0.004
average firm growth rate	8.62%	5.21%	0.10%
firm entry rate	4.09%	4.57%	0.07%
average relative size (target/acquirer)	0.316	0.337	0.014

Notes: The table reports estimation results obtained from the simulated method of moments (SMM), which chooses model parameters by matching the moments from a simulated panel of firms to the corresponding moments from the data. Panel A reports the estimated parameters. Panel B reports the simulated and actual moments. See Section 5 for the definition of moments.

Table 6: Eliminating Information Frictions

	Baseline	Short-Term Effect ($\omega = 0$) baseline $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$	Long-Term Effect ($\omega = 0$) new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$
market value	11.048	11.577	11.348
standalone value	10.392	10.392	10.304
capitalized expected gain from M&A	0.657	1.185	1.045
capitalized expected gain from M&A/market value	5.94%	10.24%	9.20%
avg. R&D intensity	5.84%	5.84%	6.43%
aggregate output	4.460	4.460	4.578
avg. merger probability	1.65%	2.30%	1.95%
consumption	2.002	2.009	2.063
avg. firm growth rate	8.62%	8.41%	8.87%

Notes: This table reports model implications in the baseline model and in a counterfactual economy in which information frictions are eliminated by setting $\omega = 0$. Column one shows the baseline results. In column two, we evaluate the short-term effect of information frictions, keeping firm innovation policy and the cross-sectional distribution as they are in the baseline equilibrium. Column three reports the long-term effect in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Market value is the model-implied value of the firm. We decompose the market value into two components: a standalone value derived by shutting down a firm's opportunity to participate in the M&A market during its lifetime, and an option value derived from the capitalized gains from all future M&A deals. R&D intensity is measured as the total expenses firms invest in innovation divided by firm assets; output is the aggregate production in the economy; avg. merger probability is the total number of mergers divided by the total number of firms in the economy; firm growth rate is the average firm sale growth.

Table 7: Alternative Benchmarks

	Baseline	$\omega = 75\% \times \omega^*$	$\omega = 50\% \times \omega^*$	$\omega = 25\% \times \omega^*$	$\omega = 0$
market value	11.048	11.119	11.181	11.274	11.348
standalone value	10.392	10.356	10.348	10.317	10.304
capitalized expected gain from M&A	0.657	0.763	0.832	0.957	1.045
capitalized expected gain from M&A/market value	5.94%	6.86%	7.45%	8.49%	9.20%
avg. R&D intensity	5.84%	5.93%	6.03%	6.33%	6.43%
aggregate output	4.460	4.488	4.492	4.550	4.578
avg. merger probability	1.65%	1.75%	1.77%	1.88%	1.95%
consumption	2.002	2.018	2.023	2.048	2.063
avg. firm growth rate	8.62%	8.70%	8.73%	8.81%	8.87%

Notes: This table reports model implications in the baseline model and in counterfactual economies in which information frictions are reduced to 75%, 50%, 25%, and 0% of its baseline value, respectively. Column one shows the baseline results. All columns report the long-term effect in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Variables are defined as in Table 6.

TABLE 8: NO EQUITY USAGE

	Baseline	Short-Term Effect baseline $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$	Long-Term Effect new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$
market value	11.048	10.796	10.803
standalone value	10.392	10.392	10.394
capitalized expected gain from M&A	0.657	0.404	0.408
capitalized expected gain from M&A/market value	5.94%	3.74%	3.78%
avg. R&D intensity	5.84%	5.84%	5.19%
aggregate output	4.460	4.460	4.323
avg. merger probability	1.65%	1.47%	1.56%
consumption	2.002	2.009	1.954
avg. firm growth rate	8.62%	8.67%	8.49%

Notes: This table reports model implications in the baseline model and in a counterfactual economy in which equity usage is prohibited. Column one shows the baseline results. In column two, we evaluate the short-term effect of shutting down equity payment, keeping firm innovation policy and the cross-sectional distribution as they are in the baseline equilibrium. Column three reports the long-term effect in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Variables are defined as in Table 6.

TABLE 9: PARAMETER ESTIMATES FOR HIGH AND LOW ANALYST FORECAST ERROR SUBSAMPLES

Low RMSE Analyst Forecast Error Subsample

ω	Λ	χ	ρ	σ	$\underline{\mu}$	ψ	γ	α
0.6543	0.0546	60.060	0.8698	0.3615	-0.0327	0.0242	3.6243	0.6359
(0.0104)	(0.0009)	(0.5496)	(0.0019)	(0.0057)	(0.0011)	(0.0007)	(0.0043)	(0.0027)

High RMSE Analyst Forecast Error Subsample

ω	Λ	χ	ρ	σ	$\underline{\mu}$	ψ	γ	α
0.7494	0.0571	45.820	0.8589	0.3886	-0.0324	0.0253	3.5850	0.6442
(0.0158)	(0.0053)	(4.0319)	(0.0071)	(0.0024)	(0.0015)	(0.0006)	(0.0387)	(0.0059)

Notes: This table reports the estimated parameters for subsamples partitioned based on information asymmetry, and the standard errors of parameter estimates are reported in parentheses. To construct the ex ante measure of information asymmetry, we obtain Earnings per Share (EPS) forecasts at the firm-analyst level from the Institutional Brokers Estimates System (I/B/E/S) database. We also obtain the realized values of firm EPS from I/B/E/S. We first measure each analyst's forecast error for a given firm-year as the difference between each analyst's forecast and the realized value of firm EPS normalized by the absolute value of realized EPS. Then, for each firm-year observation, we compute the Root Mean Square Error (RMSE) of the forecast errors. For each industry in our sample, we calculate the mean of the RMSE of analyst forecast errors. We rank industries according to this mean and split the full sample into high and low RMSE subsamples.

Table 10: Eliminating Information Frictions for High and Low Analyst Forecast Error Subsamples

	Low RMSE Subsample		High RMSE Subsample	
	Baseline	$\omega = 0$	Baseline	$\omega = 0$
		new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$		new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$
market value	10.831	11.084	11.038	11.377
standalone value	10.177	10.131	10.405	10.329
capitalized expected gain from M&A	0.654	0.953	0.633	1.049
capitalized expected gain from M&A/market value	6.04%	8.60%	5.73%	9.22%
avg. R&D intensity	5.47%	5.98%	6.09%	6.56%
aggregate output	4.318	4.415	4.516	4.622
avg. merger probability	1.56%	1.85%	1.63%	1.94%
consumption	1.940	1.991	2.020	2.077
avg. firm growth rate	8.23%	8.48%	8.57%	8.80%

Notes: This table reports model implications for the two subsamples partitioned based on information asymmetry. For each subsample estimation, we compare the baseline model and the counterfactual economy in which information frictions are eliminated by setting $\omega = 0$. Columns 1 and column 3 show the baseline results for the subsamples with low and high RMSE of analyst forecast errors, respectively. Column 2 and Column 4 report the long-term effect of eliminating information frictions in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Subsamples are partitioned as described in Table 9 and variables are defined as in Table 6.

Online Appendices:

"Acquiring Innovation Under Information Frictions" (Not for Publication)

Murat Alp Celik, Xu Tian, and Wenyu Wang *

*Contact Information: Murat Alp Celik, 229 Max Gluskin House, 150 St. George Street, Department of Economics, University of Toronto, Toronto, Ontario, M5S 3G7, Canada. Email: murat.celik@utoronto.ca; Xu Tian, B340 Amos Hall, 620 South Lumpkin Street, Department of Finance, Terry College of Business, University of Georgia, Athens, GA 30602, United States. Email: xu.tian@uga.edu; Wenyu Wang, 1309 East Tenth Street, Department of Finance, Kelley School of Business, Indiana University, Bloomington, IN 47405, United States. Email: wenywang@indiana.edu.

A Computational Appendix

In this section, we describe the algorithms used to compute an equilibrium, to simulate a panel of firms and generate model moments, and to estimate the parameter values via simulated method of moments.

A.1 Computing an Equilibrium

Given the parameter values $\beta, \omega, \Lambda, \chi, \rho, \sigma, \underline{\mu}, \psi, \gamma$, and α, we calculate the stationary general equilibrium of the model as follows:

1. Initialize a guess for the value function $V(z;\Theta)$, the innovation policy $\hat{\mu}_i(z;\Theta)$, and the time-invariant firm productivity distribution $F_s(z)$.

2. Set the value function convergence tolerance $\epsilon_V > 0$ and the firm productivity distribution convergence tolerance $\epsilon_{F_s} > 0$. Iterate the following until $\|V^{new}(z;\Theta) - V(z;\Theta)\| < \epsilon_V$ and $\|F_s^{new}(z) - F_s(z)\| < \epsilon_{F_s}$, where $\|.\|$ is defined as the sup-norm.

   (a) Value function itereration: Iterate the following until $\|V^{new}(z;\Theta) - V(z;\Theta)\| < \epsilon_V$.

      i. Given $V(z;\Theta)$, solve for the optimal acceptance/rejection decision $\hat{\mathbb{I}}_T(z_A, z_T, c, s;\Theta)$ of the target with contemporaneous productivity z_T who meets an acquirer with contemporaneous productivity z_A that makes an offer of cash c and equity s to the target. It is easy to see that whenever the merger results in a firm with positive value, for any $(z_A, z_T, s) \in \mathbb{R}_{++} \times \mathbb{R}_{++} \times [0,1]$, there exists a threshold $c^{threshold}(z_A, z_T, s)$ below which the target rejects, and above which the target accepts. Therefore, for computational efficiency, we solve for the threshold cash value $c^{threshold}$ for each value of (z_A, z_T, s) using a bisection algorithm. This implicitly defines the optimal acceptance/rejection decision as follows:

$$\hat{\mathbb{I}}_T(z_A, z_T, c, s;\Theta) = \begin{cases} 1, & \text{if } c \geq c^{threshold}(z_A, z_T, s) \\ 0, & \text{otherwise} \end{cases} \tag{20}$$

      ii. Given $V(z;\Theta)$, $\hat{\mathbb{I}}_T(z_A, z_T, c, s;\Theta)$, and $\hat{\mu}_i(z;\Theta)$, solve for the optimal cash $\hat{c}^i(z_A, \tilde{z}_T;\Theta)$ and equity $\hat{s}^i(z_A, \tilde{z}_T;\Theta)$ offered by an acquirer with contemporaneous productivity z_A, facing a target whose last period productivity is $\tilde{z}_T$, but whose contemporaneous productivity z_T is unknown to the acquirer (imperfect information scenario). Due to asymmetric information, the acquirer has to form rational expectations regarding z_T given $\tilde{z}_T$ and $\hat{\mu}_i(z;\Theta)$, and settle on the offer that maximizes its firm value, taking as given the target's response $\hat{\mathbb{I}}_T(z_A, z_T, c, s;\Theta)$. Store the implied acquirer value function under imperfect information, $V_A^i(z_A, \tilde{z}_T;\Theta)$, for later use.

      iii. Given $V(z;\Theta)$ and $\hat{\mathbb{I}}_T(z_A, z_T, c, s;\Theta)$, solve for the optimal cash $\hat{c}^p(z_A, z_T;\Theta)$ and equity $\hat{s}^p(z_A, z_T;\Theta)$ offered by an acquirer with contemporaneous productivity z_A, facing a target whose contemporaneous productivity z_T is known to the acquirer (perfect information scenario). The acquirer will settle on the offer that maximizes its firm value, taking as given the target's response $\hat{\mathbb{I}}_T(z_A, z_T, c, s; \Theta)$. Since there is no information asymmetry, the acquirer (if it chooses to make an offer that will be accepted) will make an offer that will leave the target exactly indifferent ($\hat{c}^p(z_A, z_T; \Theta) = c^{threshold}(z_A, z_T, \hat{s}^p(z_A, z_T; \Theta))$). Store the implied acquirer value function under perfect information, $V_A^p(z_A, z_T; \Theta)$, for later use.

iv. Given $V(z; \Theta)$, $F_s(z; \Theta)$, and all stored information above, solve for the optimal innovation decision $\hat{\mu}_i^{new}(z; \Theta)$ of a firm with productivity z at the beginning of the period. This involves calculating the expected value given in equation (16) associated with each possible $\mu_i \geq \underline{\mu}$, and choosing the one that maximizes the firm value. Store the implied firm value function as $V^{new}(z; \Theta)$.

v. Calculate the sup-norm $\|V^{new}(z; \Theta) - V(z; \Theta)\|$. If the norm is greater than ϵ_V, update the value function guess $V(z; \Theta) = V^{new}(z; \Theta)$ and the innovation policy guess $\hat{\mu}_i(z; \Theta) = \hat{\mu}_i^{new}(z; \Theta)$ and go back to step (i). If the norm is less than ϵ_V, continue to the next part.

(b) Stationary firm distribution: Initialize $F_s^{inner}(z) = F_s(z)$. Iterate the following until $\|F_s^{new}(z) - F_s^{inner}(z)\| < \epsilon_{F_s}$.

i. Given $F_s^{inner}(z)$ and all the information from the value function iteration regarding firm decisions, construct the transition matrix $T(z^{old}, z^{new})$ which contains the transition probability for a firm with beginning productivity z^{old} to finish the period with the new productivity z^{new}. If a firm is exogenously destroyed or acquired, the end productivity corresponds to the productivity of the new firm that replaces the exiting firm with productivity $z = \underline{z}$.

ii. Given $F_s^{inner}(z)$ and the transition matrix $T(z^{old}, z^{new})$, calculate the new guess for the firm size distribution $F_s^{new}(z)$ as $F_s^{new} = T' F_s^{inner}$.

iii. Calculate the sup-norm $\|F_s^{new}(z) - F_s^{inner}(z)\|$. If the norm is greater than ϵ_{F_s}, update the firm productivity distribution guess $F_s^{inner}(z) = F_s^{new}(z)$ and go back to step (i). If the norm is less than ϵ_{F_s}, continue to the next part.

(c) The new firm size distribution $F_s^{new}(z)$ obtained in the previous step is the stationary distribution implied by the policy functions that arise from the value function iteration step. However, for this to be an equilibrium, the firm size distribution $F^s(z)$ that was used in the value function iteration, and $F_s^{new}(z)$ must coincide. To check this, calculate the sup-norm $\|F_s^{new}(z) - F_s(z)\|$. If the norm is greater than ϵ_{F_s}, update the firm productivity distribution guess $F_s(z) = F_s^{new}(z)$ and go back to step (a). If the norm is less than ϵ_{F_s}, this means $F^s(z)$ is indeed the stationary firm size distribution associated with the stationary general equilibrium of the model, and the equilibrium is found. Stop the algorithm.

A.2 Simulating Panel Data and Replicating Regressions

In the model, the productivity z of all firms is known to the modeler, and it corresponds to the innovation stock of a company, since it is the culmination of past innovation realizations and merger activity, governed by equations (7) and (9). However, in our empirical analysis, we cannot directly observe a firm's long-term productivity, and have to construct an innovation stock by relying on patenting information. Since we target the regression coefficients we obtain in our empirical analysis to discipline the estimated parameter values, it becomes important to construct a model counterpart of the innovation stock by using the same methodology as we do in the empirical analysis, so that identical regressions can be run using model-generated panel data.[1] In this section, we describe the algorithm we use to simulate a firm panel and how we replicate the empirical regressions using this data.

Our empirical analysis spans 27 years of data from 1980 to 2006. Likewise, we choose $T = 27$ as the length for our firm panel simulation. In the data, we observe many firms that existed prior to 1980 for which we do not have past patenting information. This creates a truncation where their patent stock is started at zero despite the existence of earlier patents. To make our simulation comparable, we introduce the same truncation. This is accomplished by drawing N existing firms at $T = 1$ from the stationary firm distribution $F_s(z)$, but setting their initial innovation stock as zero.

After we draw N initial firms at $T = 1$ as described, we iterate the simulation forward using the optimal firm policy functions from the equilibrium, and using the stationary firm distribution $F_s(z)$ to determine the realized outcomes of M&A interactions. There are several stochastic draws that can occur for any firm in any given year: (1) exogenous firm exit, (2) productivity of the firm met, (3) own innovation realization, (4) met firm's innovation realization, (5) tie-breaking coin flip to become the acquirer in case the two firms have the same productivity, (6) whether there is perfect or imperfect information. We draw an i.i.d. random number from the corresponding distribution for each event $N \times T$ many times. Whenever a firm exits due to exogenous exit or endogenous acquisition, it is replaced by a new firm with $z = \underline{z}$ as in the model, which keeps the number of firms present in any year constant at N.

While conducting the simulation, we keep track of several variables: (1) the owner sequence for a line $n \in \{1, ..., N\}$, which starts at 1, and is incremented by one whenever a firm exists and is replaced by a new one, (2) the pre-innovation productivity, (3) the post-innovation productivity (before merger), (4) the end-of-period productivity (after merger), (5) an indicator that records whether the firm was acquired, (6) an indicator that records whether the firm received a merger offer, and (7) an indicator that records whether the firm acquired another firm.

Using the recorded information, we create the innovation stock for each simulated firm. The innovation stock starts at zero for all existing firms at $T = 1$, and all new firms that enter afterwards. We use the evolution of productivity that is due to innovation (i.e., not mergers or costless drift)

[1] In previous versions of the paper, we used the productivity z directly as the model counterpart of the innovation stock. Although the two are very highly positively correlated, they are not exactly the same, which the explanation below clarifies. We thank our anonymous referee for their suggestion.

to back out the innovation realization for each firm in each period. Recall the law of motion for productivity given in equation (7). Using this equation we can calculate the log productivity gain that is due to innovation as:

$$(\mu_{it} - \underline{\mu}) + \sigma_\epsilon \epsilon_{it+1} = \log(z_{it+1}) - \rho \log(z_{it}) - \underline{\mu} \tag{21}$$

where $\underline{\mu}$ is subtracted since only $(\mu_{it} - \underline{\mu})$ of the drift is due to the R&D expenditures of the firm, $X(\mu_{it}, z_{it}) = \chi(\mu_{it} - \underline{\mu})^2 z_{it}$. To get back at the implied patent citations associated with the innovation, we assume a linear relationship between the productivity gain due to the firm's innovation, and the citations the associated patents receive.[2] Given the innovation flows for each firm in each year, it becomes possible to construct the innovation stock variable via forward iteration. Mimicking what we do in the empirical analysis, we construct the stock using the perpetual inventory method with a depreciation rate of 6%.

With the innovation stock constructed in the same way as in the empirical analysis, we run the three regressions that have the patent stock variables as regressors using the simulated firm panel. To make the model-generated innovation stock variable comparable to that in the data, we normalize and rescale it such that it has the same mean and standard deviation as the corresponding variable in each regression. In our estimation, we set $N = 20,000$, which results in a panel with $N \times T = 540,000$ observations. This results in roughly $500,000$ observations for the model counterpart of the regression outlined in equation (1), compared to $57,885$ observations in the data. The observation count is roughly $10,000$ for the model counterparts of regressions given in equations (2) and (3), compared to 590 and 663 observations in the data, respectively. For robustness, we verify that the associated model moments do not change meaningfully even if we increase the number of simulated firms N to $1,000,000$, which results in a final sample of $27,000,000$ observations.

A.3 SMM Estimation

We use the simulated method of moments (SMM) to estimate our model. We first calibrate the parameters that are standard in the literature. The first three parameters are the subjective discount rate β, capital depreciation rate δ, and the production function concavity on capital κ. We set $\beta = 0.9615$, consistent with a real interest rate of $r = 0.04$, $\delta = 0.069$ is taken from U.S. NIPA, and $\kappa = 0.85$, consistent with an average markup of 18%.

We are left with 9 parameters to be estimated: $\omega, \Lambda, \chi, \rho, \sigma, \underline{\mu}, \psi, \gamma$ and α. We use 14 moments to identify the remaining 9 model parameters. Our identification strategy ensures that there is a unique parameter vector that makes the model fit the data as closely as possible. Since we estimate these

[2] This assumption of a positive linear relationship between long-term productivity increase and metrics of patent quality is used in the vast majority of the endogenous growth and innovation literatures. Still, one might be concerned that the true relationship might be concave or convex instead of linear. Using our data, we demonstrate that the observed relationship between annual changes in measures of firm productivity and our patent quality metrics is almost perfectly linear, in line with our assumption. Results are available upon request.

parameters in one big SMM system, we essentially allow each moment to respond to all parameters in estimation.

SMM proceeds in the following way: For an arbitrary value of parameter vector $\theta = \{\omega, \Lambda, \chi, \rho, \sigma, \underline{\mu}, \psi, \gamma, \alpha\}$, the dynamic problem is solved, the policy functions are generated, and a firm panel is simulated as described in the two preceding sections. Let x_{it} be the actual data vector, $i \in \{1, ..., N\}$, $t \in \{1, ..., T\}$, and let $y_{it}(b)$ be the simulated vector corresponding to $i \in \{1, ..., N\}$, and $t \in \{1, ..., T\}$. The simulated data vector, $y_{it}(\theta)$, depends on a vector of structural parameters, θ. Define the moment conditions as:

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[h(x_{it}) - h(y_{it}(\theta))\right] \equiv \Psi^A - \Psi^S(\theta) \tag{22}$$

where $h(y_{it}(\theta))$ is a vector of simulated moments and $h(x_{it})$ is the actual data moments. $\Psi^A = \frac{1}{NT}\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} h(x_{it})$, $\Psi^S(\theta) = \frac{1}{NT}\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} h(y_{it}(\theta))$

The simulated moments estimator is defined as the solution to the minimization of:

$$\hat{\theta} = \arg\min_{\theta} \left[\Psi^A - \Psi^S(\theta)\right]' \hat{W} \left[\Psi^A - \Psi^S(\theta)\right] \tag{23}$$

in which $\hat{W}$ is a positive definite matrix that converges in probability to a deterministic positive definite matrix W. It is constructed by calculating the inverse of the variance-covariance matrix of the data moments. Define Ω as the variance-covariance matrix of the data moments Ψ^A. Lee and Ingram (2010) show that under the estimating null, the variance-covariance of the simulated moments $\Psi^S(\theta)$ is equal to $\frac{1}{K}\Omega$, where K is the number of simulated panels, i.e. the ratio of N over the corresponding number of firms in the data. Since $\Psi^A$and $\Psi^S(\theta)$ are independent by construction, $\hat{W} = \left[(1 + \frac{1}{K})\Omega\right]^{-1}$. Ω is calculated using influence function method following Erickson and Whited (2002).

We use a simulated annealing algorithm for minimizing the objective function. This starts with a predefined first and second guess. For the third guess onward, it takes the best prior guess and randomizes from this to generate a new set of parameter guesses. That is, it takes the best-fit parameters and randomly "jumps off" from this point for its next guess. Over time the algorithm "cools", so that the variance of the parameter jumps falls, allowing the estimator to fine-tune its parameter estimates around the global best fit. We restart the program with different initial conditions to ensure the estimator converges to the global minimum. The simulated annealing algorithm is extremely slow, which restricts the size of the parameter space that can be estimated. Nevertheless, we use this because it is robust to the presence of local minima and discontinuities in the objective function across the parameter space.

The simulated moments are asymptotically normal for fixed K. Denote $g(\theta) \equiv \Psi^A - \Psi^S(\theta)$. The asymptotic distribution of θ is given by:

$$\sqrt{n}(\theta - \hat{\theta}) \xrightarrow{d} N(0, avar(\hat{\theta})) \tag{24}$$

in which

$$avar(\hat{\theta}) = (1+\frac{1}{K}) \left[\frac{\partial g}{\partial \theta} W \frac{\partial g}{\partial \theta'}\right]^{-1} \left[\frac{\partial g}{\partial \theta} W \Omega W \frac{\partial g}{\partial \theta'}\right] \left[\frac{\partial g}{\partial \theta} W \frac{\partial g}{\partial \theta'}\right]^{-1} \tag{25}$$

in which Ω is the probability limit of a consistent estimator of the covariance matrix. We calculate the estimate of this covariance matrix using influence function of the moment vector clustered at firm level following Erickson and Whited (2002).

B Theory Appendix

B.1 Two-Sided Information Asymmetry

In the baseline model, we assume that the target firm always learns the acquirer's contemporaneous productivity through contact, whereas the acquirer firm's due diligence does not always reveal the target's contemporaneous productivity. Therefore, the information asymmetry is one-sided, rather than two-sided. This assumption eliminates the complications that arise from two-sided information asymmetry, such as multiplicity of equilibria, which requires taking a stance on equilibrium selection, and increases the computational burden. To compensate for the lack of two-sided information asymmetry, the baseline model includes a reduction in the value of the merged firm, $g(c, s, z_m; \Theta')$, which captures the negative market reaction to equity usage in a reduced form way. The market valuation of the firm is decreasing in the share of the transaction fee paid in the form of equity instead of cash.

In this section, we lay out a version of the stage game with two-sided information asymmetry. This serves two main purposes: First, the exercise exhibits how multiple equilibria can arise. Equilibrium multiplicity is a well-recognized outcome in the two-sided information asymmetry literature, and purification methods remain elusive. The same outcome is encountered in our economic setting. Second, we show how much the target firms would discount the acquirer's equity offer, and how this discounted amount would vary with the fundamentals such as the dispersion of acquirers' productivity, and the beliefs regarding the probability of meeting a low productivity acquirer. Importantly, we show that the optimal discounting of equity is the same across all pooling equilibria, and consistent with the $g(c, s, z_m; \Theta')$ function. This relationship provides the motivation for the use of the reduced-form negative market reaction to equity usage, $g(c, s, z_m; \Theta')$, in the baseline model: regardless of which pooling equilibrium is chosen, the optimal discounting of equity is correctly captured by the reduced-form functional form, and its scale parameter that measures the severity can then be estimated using the data.

B.1.1 Environment and Information Structure

Consider the problem of an acquirer and a target who have already met each other under two-sided information asymmetry. Both parties know the distributions from which the true productivities are drawn, but the realized productivities are the private information of the firms alone.[3] As in the baseline model, the acquirer can make an offer that consists of cash and equity. After receiving an offer, the target can accept it, in which case the two firms merge; or reject the offer, conditional upon which the two firms remain stand-alone forever.

[3] In other words, we are considering the case where due diligence has failed to reveal any information regarding the opposing party for both sides, which is the only scenario with two-sided information asymmetry. The case where due diligence succeeds for both parties corresponds to Scenario 2 in Section 3.3.4. The case where due diligence succeeds only for the target firm corresponds to Scenario 1 in Section 3.3.4. The case where due diligence succeeds only for the acquirer firm is a special case of the two-sided information asymmetry case we currently consider, with a degenerate distribution for z_T.

Different from the baseline model, the true productivity of the acquirer firm, z_A, is unknown to the target. For expositional simplicity, assume that the true productivity of the acquirer is drawn from the discrete set $z_A \in \{z_A^H, z_A^L\}$ with $z_A^H > z_A^L = \iota z_A^H$ and $\iota \in (0,1)$.[4] Denote the probabilities of the two outcomes as $p = Pr(z_A = z_A^H)$ and $1-p = Pr(z_A = z_A^L)$ respectively, with $p \in (0,1)$. Linking this to the baseline model, z_A^H can be thought of as a successful innovation realization which delivers a high contemporaneous productivity compared to its expected value ($z' > \mathbb{E}[z'|z,\mu_i;\Theta]$), whereas z_A^L stands for a lackluster innovation realization with a worse-than-expected outcome ($z' < \mathbb{E}[z'|z,\mu_i;\Theta]$). Since the target does not know the true type of the acquirer z_A, it will try to infer the true type conditional on the observed merger offer. This turns the described interaction into a signalling game, and given the two types, the described game can have a separating or a pooling Perfect Bayesian Equilibrium.[5] All details regarding the production and the merger technologies are the same as in the baseline model. We discuss the decision problems of the two agents next.

B.1.2 Target's problem:

Consider the problem of the target firm with productivity z_T which receives a merger offer of cash and equity, (c,s), from the acquirer. Denote the probability the target attaches to the acquirer being of high type z_A^H conditional on receiving an offer as $q(c,s) = Pr(z_A = z_A^H|c,s)$ with $q(c,s) \in [0,1], \forall c,s$, which fully defines its beliefs. The target firm must decide whether to accept or reject the offer by comparing its continuation value and the offer:

$$\max_{\mathbb{I}_T \in \{0,1\}} \left\{ \mathbb{I}_T \left((1-s)c + s\mathbb{E}\left[R\pi\gamma z_A^\alpha z_T^{1-\alpha} | c,s \right] \right) + (1-\mathbb{I}_T) R\pi z_T \right\} \tag{26}$$

where $R \equiv \sum_{t=0}^{\infty} \left(\frac{1-\psi}{1+r}\right)^t = \frac{1+r}{r+\psi}$ discounts all future profits to today and $\pi = \zeta(\frac{\kappa}{r+\delta})^{\frac{\kappa}{1-\kappa}}$ is the time-invariant constant that transforms firm productivity to static profit flows from production, derived in the baseline model. The first term represents the offer value perceived by the target and the second term is the continuation value if it rejects the offer. We can explicitly derive the equation that pins down the offers (c,s) which leave the target indifferent between accepting or rejecting the offer as follows:

$$(1-s)c + sR\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha} \left(q(c,s) + (1-q(c,s))\iota^\alpha \right) = R\pi z_T \tag{27}$$

In the case of separation, we have:

$$(1-s)c + sR\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha} = R\pi z_T, \quad \text{if } q(c,s) = 1 \tag{28}$$

$$(1-s)c + sR\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha}\iota^\alpha = R\pi z_T, \quad \text{if } q(c,s) = 0 \tag{29}$$

[4] Naturally, this assumption can be relaxed to allow for an arbitrary distribution over z_A, and all the insights would carry through unchanged.

[5] And semi-pooling equilibria if we have more than two types.

In the case of pooling with $q(c,s) = p$, we have:

$$(1-s)c + sR\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha}\left(p + (1-p)\iota^\alpha\right) = R\pi z_T \tag{30}$$

Suppose that the target accepts the offer if it is indifferent. We denote the acceptance/rejection decision that solves this problem as $\hat{\mathbb{I}}_T(z_T, c, s)$.

B.1.3 Acquirer's problem:

Now consider the problem of the acquirer firm with true productivity $z_A \in \{z_A^H, z_A^L\}$. It has rational expectations over the true productivity of the target, z_T.[6] The acquirer takes the beliefs of the target $q(c,s), \forall c, s$ as given. The acquirer firm must decide on the merger payment (c,s) that will be offered to the target:

$$\max_{c\geq 0, s\in[0,1]} \quad \left\{\mathbb{E}\left[\mathbb{I}_T\left((1-s)\left(-c + R\pi\gamma z_A^\alpha z_T^{1-\alpha}\right)\right) + (1-\mathbb{I}_T)R\pi z_A | q(c,s)\right]\right\} \tag{31}$$

$$\mathbb{I}_T = \hat{\mathbb{I}}_T(z_T, c, s) \tag{32}$$

The first term denotes the value to the acquirer firm if the offer is accepted ($\mathbb{I}_T = 1$). If the offer is accepted, the acquirer retains $1-s$ fraction of the combined firm and pays c to the target in cash. The second term denotes the value if the offer is rejected, in which case the acquirer remains stand-alone. The acquirer chooses the offer (c,s) optimally, having rational expectations over the target's true productivity, and given the acceptance/rejection decision, $\hat{\mathbb{I}}_T(z_T, c, s)$, and the beliefs, $q(c,s)$. We denote the cash and equity components of the optimal offer as $\hat{c}(z_A)$ and $\hat{s}(z_A)$, respectively.

B.1.4 Pooling equilibria:

Given the decision problems of the target and the acquirer, we can characterize the set of pooling equilibria that can be supported as a Perfect Bayesian Equilibrium in pure strategies under some beliefs $q(c,s)$. Consider the beliefs $q^*(c,s;c^*,s^*)$ defined as follows:

$$q^*(c,s;c^*,s^*) = \begin{cases} p, & \text{if } c = c^* \text{ and } s = s^* \\ 0, & \text{otherwise} \end{cases} \tag{33}$$

According to these beliefs, the target attaches the unconditional probability p to the acquirer being the high-type when the offer received is exactly (c^*, s^*), and attaches a probability zero to the acquirer being the high-type for any other offer. The described beliefs would be consistent with a pooling equilibrium where both types of acquirers make the offer (c^*, s^*). Under these beliefs, we can characterize the conditions necessary for the existence of such a pooling equilibrium.

[6] In the baseline model, these rational expectations are formed based on the target's productivity in the previous period, $\tilde{z}_T$, and the innovation policy $\mu_i(\tilde{z}_T)$ which are both public information.

To begin, let's consider how targets of varying productivities react to offers given some beliefs. First, note that for any z^T, the acceptance set under $q(c,s) = p$ is a proper superset of the acceptance set under $q(c,s) = 0$. This can be gleaned from Equations 29 and 30. In both cases, the target accepts an only-cash offer of $R\pi z_T$ or above. However, the target accepts equity at a discount $\iota^\alpha < p + (1-p)\iota^\alpha$ in the latter case compared to the prior case. Therefore, the same offer (c,s) is more attractive to the target when $q(c,s) = p$ instead of $q(c,s) = 0$. Second, note that for any z_T^H and z_T^L with $z_T^H > z_T^L$, and under any constant belief $q(c,s) = q$, the acceptance set of z_T^H is a proper subset of z_T^L. This can be seen from Equation 27. This result is also intuitive: targets with higher productivity are more selective when it comes to accepting an offer, since their stand-alone value is higher.

Now, consider the problem of the acquirers with low type, z_A^L, and high type, z_A^H. Consider a pooling equilibrium with the offer (c^*, s^*) made by both types, with beliefs $q(c,s;c^*,s^*)$ specified as before. For this equilibrium to be rational for the high type, we need:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{I}_T(1-s^*)\left(-c^* + R\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha}\right) + (1-\mathbb{I}_T)R\pi z_A^H | q(c,s) = p\right] \geq \\
&\max_{c\geq 0, s\in[0,1]}\{\mathbb{E}\left[\mathbb{I}_T(1-s)\left(-c + R\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha}\right) + (1-\mathbb{I}_T)R\pi z_A^H | q(c,s) = 0\right]\}
\end{aligned} \tag{34}$$

In other words, the prescribed strategy (c^*, s^*) must deliver a utility to the high type above what can be achieved by any other strategy under the constant beliefs where targets always assume the acquirer to be the low type. Given the harsher discounting of equity under $q(c,s) = 0$, ($\iota^\alpha < p + (1-p)\iota^\alpha$ as discussed before), a region of such (c^*, s^*) with positive Lebesgue measure can be found. For this equilibrium to be rational for the low type, we need:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{I}_T(1-s^*)\left(-c^* + R\pi\gamma(z_A^L)^\alpha z_T^{1-\alpha}\right) + (1-\mathbb{I}_T)R\pi z_A^L | q(c,s) = p\right] \geq \\
&\max_{c\geq 0, s\in[0,1]}\{\mathbb{E}\left[\mathbb{I}_T(1-s)\left(-c + R\pi\gamma(z_A^L)^\alpha z_T^{1-\alpha}\right) + (1-\mathbb{I}_T)R\pi z_A^L | q(c,s) = 0\right]\}
\end{aligned} \tag{35}$$

Intersection of the set that satisfies (35) with the set that satisfies (34) is non-empty for most relevant parameter values, but not always. When the sets have a non-empty intersection, the intersection generically has positive Lebesgue measure, which implies the existence of not just one pooling equilibrium, but a two-dimensional continuum of pooling equilibria which can all be maintained as a Perfect Bayesian Equilibrium in pure strategies (i.e., uncountably many pooling equilibria exist.) When the intersection is empty, we have a separating equilibrium instead. This only happens when the low type is so unproductive compared to the target that they know the expected gains from the merger are not sufficient to raise c^* despite the huge over-valuation of their shares thanks to mimicking the high type.

The fringe cases where we have a separating equilibrium instead of pooling equilibria are not of great interest, given that the information friction problem resolves itself, and we are back to one-sided information asymmetry in effect. The case of pooling equilibrium, on the other hand, is interesting, because the existence of one generically implies the existence of a continuum of pooling equilibria. This is a well-known outcome in the two-sided information asymmetry literature, which is summarized in great detail by Ausubel, Cramton, and Deneckere (2002) in the third volume of the Handbook of Game Theory. Reviewing the sequential bargaining models with two-sided incomplete information, Ausubel, Cramton, and Deneckere (2002) mention that "[w]hile the (upper) boundary of the set of all sequential equilibria is known, little exists in the way of results refining the set of sequential equilibrium outcomes." Even with alternating offers, it is difficult to rule out multiplicity of equilibria, and the purification criteria proposed in the literature are found to exhibit undesirable properties. Thus, equilibrium selection criteria become necessary even when one studies a two-sided information asymmetry problem in isolation, before one embeds the stage game into a dynamic general equilibrium model with heterogeneous firms, endogenous productivity growth, and search frictions. Fortunately, the optimal equity discounting undertaken by the targets is found to be the same for all possible pooling equilibria, which is shown in the next section. Due to this shared property across all equilibria, it becomes possible to construct a reduced-form function $g(c, s, z, m; \Theta')$ which captures the negative market reaction to equity resulting from the acquirer-side information asymmetry that is valid regardless of the equilibrium selection criterion.

B.1.5 Optimal equity discounting and its properties

When considering the solution of the target's decision problem, it was observed that the equity payment by the high type acquirer was accepted at a discount by the target in a pooling equilibrium because of the risk of over-payment to an acquirer of low type. We can directly calculate how much the targets discount equity usage of the acquirer, and how the optimal discounting is related to the fundamental parameters that determine the severity of the information asymmetry regarding the acquirer's true type.

Consider any pooling equilibrium with the offer (c^*, s^*) made by both acquirer types, with beliefs $q(c, s; c^*, s^*)$ specified as before. A target with true productivity z_T evaluates the expected value of this offer as:

$$(1 - s^*)c^* + s^* R\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha} \left(p + (1-p)\iota^\alpha\right) \tag{36}$$

If the acquirer is of the high type z_A^H, the realized (i.e. true) value of the offer is:

$$(1 - s^*)c^* + s^* R\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha} \tag{37}$$

The difference between the two gives us the amount by which the offer is discounted, which is:

$$\begin{aligned} & s^* R\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha}(1 - (p + (1-p)\iota^\alpha)) \\ = \quad & s^* R\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha}(1-p)(1-\iota^\alpha) \end{aligned} \tag{38}$$

This tractable closed-form formulation of the target's discounting exhibits three important properties:

1. The discounted amount is linearly increasing in the unconditional probability of low type

acquirers, $(1-p)$.

2. The discounted amount is decreasing in the relative productivity of the low type acquirers, ι.

3. The discounted amount is linearly increasing in the shares offered, s^*.

The first two properties show that when the information asymmetry on the acquirer side is more severe, the targets discount the usage of equity more. If low type acquirers are more likely to be encountered, or if their productivity is much lower compared to the high type, the targets rationally revise the expected value of equity payment downwards. Second, the total discount amount is linearly increasing in the use of shares, which means the expected value of the combined firm is decreasing in equity usage.[7]

The reduced-form function $g(c,s,z_m;\Theta')$ we use in the baseline model is consistent with these observations. Recall the functional form we used:

$$\begin{aligned} g(c,s,z_m;\Theta') &= \Lambda E(c,s,z_m;\Theta')z_m \\ &= \Lambda \left[\frac{s\left(\pi z_m + \frac{1-\psi}{1+r}V(z_m;\Theta')\right)}{c+s\left(-c+\pi z_m+\frac{1-\psi}{1+r}V(z_m;\Theta')\right)} \right] z_m \\ &= \Lambda s\left(\pi z_m + \frac{1-\psi}{1+r}V(z_m;\Theta')\right) \left[\frac{z_m}{c+s\left(-c+\pi z_m+\frac{1-\psi}{1+r}V(z_m;\Theta')\right)} \right] \end{aligned} \quad (39)$$

Holding the last term constant, the negative market reaction to equity usage this function imposes is linear in the shares used, s, and the total value of the merged firm without any discounting, $\pi z_m + \frac{1-\psi}{1+r}V(z_m;\Theta')$.[8] There is a clear mapping between the $g(c,s,z_m;\Theta')$ function and Equation 38. Comparing this to Equation 38, s corresponds to s^*, and $\pi z_m + \frac{1-\psi}{1+r}V(z_m;\Theta')$ corresponds to $R\pi\gamma(z_A^H)^\alpha z_T^{1-\alpha}$, which is the non-discounted value of the merged firm if the acquirer is of high-type. The final term $(1-p)(1-\iota^\alpha)$, as discussed before, is increasing in the degree of the severity of the information asymmetry on the acquirer side. The scale parameter Λ in the $g(c,s,z_m;\Theta')$ function is estimated to capture this severity, which we discipline by targeting the empirically-observed relationship between equity usage and announcement returns. In the two-sided information asymmetry model, there is no discounting if $(1-p)(1-\iota^\alpha)=0$. In the baseline model, there is no discounting if $\Lambda=0$.

To summarize, although incorporating two-sided information asymmetry to the baseline model is not desirable due to the high computational burden and the need for explicit equilibrium selection criteria, the $g(c,s,z_m;\Theta')$ function is able to incorporate the trade-offs we observe when we solve

[7] It is important to note that the movement in s^* is a movement *across* equilibria within the continuum of pooling equilibria defined above. That is, this is a result that holds with the same fundamental parameter values, but across different pooling equilibria, which are all feasible.

[8] The last term is a merger-specific term that captures the relative value of the merged firm compared to that of the target. There is little cross-sectional variation given that the value of the merged firm and the total payment to the target are positively associated.

the stage game with two-sided information asymmetry in a reduced form way, the impact of which we discipline using the real-world reaction of announcement returns to equity usage.[9]

C Extended Model with Target-Side Lack of Information

In the baseline model, we assume that the target firm is always informed about its own contemporaneous productivity. However, in reality, there can be cases where the target firm is also unable to assess the value of its own recent innovotion. In such a scenario, the target firm would have to make a decision to accept or reject a bid under uncertainty, and the acquirer must make an offer taking this uncertainty into account. In this section, we extend our model to allow for such scenarios, pin down the probability of such a scenario using information on patent renewals, and show that our results remain largely robust under a large range of parametrizations.

C.1 Extending the Model

In the baseline model, there were two scenarios to consider: the case with imperfect information, where the target knows its own contemporaneous productivity, but the acquirer does not have access to this information; and the case with perfect information, where both the target and the acquirer know the contemporaneous productivity of both firms. On top of these two scenarios, we add a third one: a case where both the acquirer and the target do not know the contemporaneous productivity of the target firm. Therefore, both firms must use rational expectations to formulate their optimal strategies. Let $\Upsilon \in [0,1]$ denote the probability of this new scenario. This results in the following modified timeline of events:

1. Firms choose their innovation policy μ_i and incur the R&D cost $X(\mu_i, z)$.
2. Matches in the merger market are realized.
3. Two matched firms choose their roles as the acquirer or as the target.
4. Innovation outcomes are realized, and the contemporaneous productivity of the acquirer and the target are updated.
5. With probability $(1-\omega)(1-\Upsilon)$, the target's contemporaneous productivity is revealed to the acquirer, and with probability $\omega(1-\Upsilon)$, it remains the private information of the target. With probability Υ, neither the acquirer nor the target know the target's contemporaneous productivity.
6. The acquirer firm makes an offer to the target, specifying the method of payment.

[9] Notice that our baseline model's resulting equilibrium does correspond to one amongst the many multiple equilibria that arise from the extended model. What we do is to allow the acquirer to select amongst the multiple equilibria that maximize its value subject to the optimal equity discounting due to the pooling equilibrium, instead of imposing an arbitrary equilibrium selection criterion.

7. The target decides whether to accept or decline the offer.

The decisions to merge for the target and the acquirer firms must be updated to incorporate this new third scenario. Consequently, the innovation decision must also be revisited. We provide the details below.

C.1.1 Target Firm's Problem Under Lack of Information

Consider the problem of a target firm which had the productivity $\tilde{z}_T$ in the previous period, but is not informed about its own contemporaneous productivity z_T. It receives a merger offer of cash and equity, (c, s), from an acquirer firm with technology z_A. It has chosen the innovation policy μ_i at the beginning of the period. We denote the value function of the target under this scenario as $V_T^l(z_A, \tilde{z}_T, c, s, \mu_i; \Theta)$, where the superscript l stands for *lack of information*. The target firm decides whether to accept or reject the offer by comparing its continuation value and the offer price:

$$\begin{aligned} V_T^l(z_A, \tilde{z}_T, c, s, \mu_i; \Theta) = \max_{\mathbb{I}_T \in \{0,1\}} \quad \Bigg\{ \quad & \mathbb{I}_T \left(c + s\mathbb{E}\left[-c + \pi z_m - g(c, s, z_m; \Theta') + \frac{1-\psi}{1+r} V(z_m; \Theta') \,\middle|\, \tilde{z}_T, z_A, \mu_i; \Theta \right] \right) \\ & + (1 - \mathbb{I}_T)\mathbb{E}\left[\pi z_T + \frac{1-\psi}{1+r} V(z_T; \Theta') \,\middle|\, \tilde{z}_T, \mu_i; \Theta \right] \Bigg\} \end{aligned} \tag{40}$$

The first term represents the expected offer value perceived by the target and the second term is the target's continuation value if it rejects the offer. Different from the scenarios considered in the baseline model, under this scenario, the target has to form rational expectations over its own contemporaneous productivity z_T (and consequently, the productivity of the merged firm z_m) based on its productivity in the previous period, $\tilde{z}_T$, and the innovation policy it chose, μ_i. We denote the acceptance/rejection decision that solves the above maximization problem as $\hat{\mathbb{I}}_T^l(z_A, \tilde{z}_T, c, s, \mu_i; \Theta)$ accordingly.

C.1.2 Acquiring Firm's Problem Under Lack of Information

Since the target's contemporaneous productivity is not revealed to either firm, the acquirer only observes the target's last period productivity $\tilde{z}_T$ and would try to form a rational expectation of the target's contemporaneous productivity z_T based on its information set. We denote the value function of the acquirer under this scenario as $V_A^l(z_A, \tilde{z}_T; \Theta)$, where the superscript l stands for *lack of information*. The maximization problem is set up as:

$$\begin{aligned} V_A^l(z_A, \tilde{z}_T; \Theta) = \max_{c \geq 0, s \in [0,1]} \quad \Bigg\{ \quad & \mathbb{E}\Bigg[\mathbb{I}_T \left((1-s) \left(-c + \pi z_m - g(c, s, z_m; \Theta') + \frac{1-\psi}{1+r} V(z_m; \Theta') \right) \right) \\ & + (1 - \mathbb{I}_T) \left(\pi z_A + \frac{1-\psi}{1+r} V(z_A; \Theta') \right) \Bigg| \tilde{z}_T, z_A; \Theta \Bigg] \Bigg\} \end{aligned} \tag{41}$$

$$\mathbb{I}_T = \hat{\mathbb{I}}_T^l(z_A, \tilde{z}_T, c, s, \mu_i; \Theta) \tag{42}$$

The decision problem of the acquiring firm is similar to the one under imperfect information (scenario 1), with the only difference being the decision rule of the target firm, which is now $\hat{\mathbb{I}}^l_T(z_A, \tilde{z}_T, c, s, \mu_i; \Theta)$ instead of $\hat{\mathbb{I}}_T(z_A, \tilde{z}_T, c, s; \Theta)$. In other words, the acquiring firm knows that the target does not know its own contemporaneous productivity z_T either, and formulates its own offer accordingly. We denote the cash and equity components of the optimal offer as $\hat{c}^l(z_A, \tilde{z}_T; \Theta)$ and $\hat{s}^l(z_A, \tilde{z}_T; \Theta)$ under this lack of information scenario.

C.1.3 Firm Innovation Decision in the Extended Model with Lack of Information

Given the expected merger gains described in the main text and above, we can now characterize firms' optimal innovation decisions. At the beginning of each period, firms choose their innovation policy μ_i to maximize the expected firm value:

$$
\begin{aligned}
V(z;\Theta) \;=\;& \max_{\mu_i \geq \underline{\mu}} \Bigg\{ \int \Bigg[\mathbb{I}_{acq}(z, z_o) \Big((1-\Upsilon)\omega \mathbb{E}\left[V^i_A(z', z_o; \Theta) \Big| z, \mu_i; \Theta \right] \\
+\;& (1-\Upsilon)(1-\omega)\mathbb{E}\left[V^p_A(z', z'_o; \Theta) \big| z, \mu_i; \Theta \right] \\
+\;& \Upsilon \mathbb{E}\left[V^l_A(z', z_o; \Theta) \Big| z, \mu_i; \Theta \right] \Big) \\
+\;& \big(1 - \mathbb{I}_{acq}(z, z_o)\big) \Big((1-\Upsilon)\omega \mathbb{E}\left[V_T(z'_o, z', \hat{c}^i(z'_o, z), \hat{s}^i(z'_o, z); \Theta) \Big| z, \mu_i; \Theta \right] \\
+\;& (1-\Upsilon)(1-\omega)\mathbb{E}\left[V_T(z'_o, z', \hat{c}^p(z'_o, z'), \hat{s}^p(z'_o, z'); \Theta) \big| z, \mu_i; \Theta \right] \\
+\;& \Upsilon \mathbb{E}\left[V^l_T(z'_o, z, \hat{c}^l(z'_o, z), \hat{s}^l(z'_o, z), \mu_i; \Theta) \Big| z, \mu_i; \Theta \right] \Big) \Bigg] dF_s(z_o) \\
-\;& X(\mu_i, z) \Bigg\}
\end{aligned}
\tag{43}
$$

The two new terms with the factor Υ account for the expected firm value under the lack of information scenario when the firm is an acquirer (the third term), and a target (the sixth term). As in the baseline model, denote the optimal innovation policy function that solves the problem as $\hat{\mu}_i(z; \Theta)$.

C.2 Quantitative Results with the Extended Model with Lack of Information

The extended model has an additional parameter Υ which denotes the probability that both the target and the acquirer lack information on the target's contemporaneous productivity. In other words, even the target firm is ill-informed about the quality of its own innovation, indicating a scenario where there is uncertainty, but no asymmerty in information. Its value needs to be pinned down to simulate the model. The micro-data on patent renewals provided by the USPTO offer a unique opportunity to potentially uncover the frequency of such an event.

In the United States, inventing firms have to incur large costs to file their patents in the form of application and patent attorney fees on top of the research and development costs to come up with the invention in the first place.[10] Therefore, they only apply for a patent if the expected value of patenting the innovation is above the application costs. Once a patent is granted, the firms must also pay a patent maintenance fee in the fourth, eighth, and twelfth years after the grant date to maintain the patent. Unlike the R&D and application costs, patent maintenance fees are quite trivial.[11] However, we observe that some firms choose not to renew their patents despite the trivial maintenance fees. When such an event is observed, we can infer that the firm revised its own valuation of the patent, and decided that its value is below this trivial maintenance fee. The administrative data on patent renewals can therefore inform us regarding the frequency of events where a patent turns out to be a "dud" – patents that were initially thought to be worthwhile, but turned out to be of insignificant value ex-post.

The frequency of patents that are not renewed in their fourth year is 9.15% for the patents granted to the sample of firms we study. One possible parametrization is to set $\Upsilon = 9.15\%$. For robustness, we consider low, medium, high, and extreme lack of information scenarios, which correspond to $\Upsilon = \frac{9.15}{2} = 4.57\%$, $\Upsilon = 9.15$, $\Upsilon = 9.15 \times 2 = 18.3\%$, and $\Upsilon = 9.15 \times 3 = 27.44\%$ respectively. We use the extended model with lack of information under the three parametrizations to reconduct our quantitative experiments in Section 6, and compare the results with those obtained in the benchmark model in the main text.

The results of the experiments can be found in Tables E5-E8. A higher value for Υ reduces the likelihood of the imperfect information case (probability: $(1 - \Upsilon)\omega$) as well as the perfect information case (probability: $(1 - \Upsilon)(1 - \omega)$). Consequently, we observe that the quantitative importance of information frictions is diminished. However, this reduction in the importance of information frictions remains quite limited. In the benchmark model (which is equivalent to setting $\Upsilon = 0$ in the extended model), we had found that removing information frictions would increase the capitalized expected gain from M&A by 59.1% of its value. Under the low, medium, high, and extreme lack of information parametrizations, this increase is calculated to be 55.5%, 52.0%, 45.3%, and 35.8% respectively. These results show that information frictions in M&A are still quite important at 35.8% even under the extreme lack of information parametrization which is the most conservative one.

If we consider the impact on consumer welfare, the consumption-equivalent welfare gain of removing information frictions was 3.02% in the benchmark model. Under the low, medium, high, and extreme lack of information parametrizations, the welfare gain becomes 2.57%, 2.46%, 2.24%, and 1.84% respectively. In other words, the welfare gain from removing information frictions is reduced by only one third of its benchmark value even under the extreme scenario. We take these results to indicate that our baseline results remain largely robust to considerations regarding target-side lack of information on innovation quality.

[10] For utility patents, patent attorney fees can go up to $40,000 and beyond for complex patents.

[11] The first payment due at 3.5 years is $2,000 for firms with more than 500 employees, $1,000 for firms with less than 500 employees, and $500 dollars for micro entities as defined by the USPTO.

D Robustness of the Inverted-U Relationship

In the baseline empirical analysis, we have sought to identify an inverted-U relationship between the probability of being acquired and firm innovation stock using the standard approach found in numerous economic studies investigating such non-linear relationships. This involves running regressions with linear and quadratic terms, establishing the significance of their coefficients, and showing that the extremum lies within the data range. Lind and Mehlum (2010) develop a hypothesis test for the existence of U and inverted-U-shape relationships.[12] To further establish the robustness of our results, we conduct the hypothesis test proposed in Lind and Mehlum (2010) for all specifications in Table 2, where the null hypothesis is the lack of an inverted-U relationship. This involves testing whether or not the slope of the curve is positive at the start and negative at the end of the interval of the variable of interest. Correspondingly, Table E4 reports the t and p-values at the lower and upper bounds of the interval of the explanatory variable. The null hypothesis is firmly rejected in all specifications. The inverted-U relationships that we have identified pass the formal test of existence, with p-values below 1% in the vast majority of cases.

E Additional Tables

[12] Arcand, Berkes, and Panizza (2015), Rodrik (2016), Bazzi, Gaduh, Rothenberg, and Wong (2016), Kesavan, Staats, and Gilland (2014), Tan and Netessine (2014), and Batt and Terwiesch (2017) among others use the test proposed in Lind and Mehlum (2010) to establish the existence of U- and inverted-U-shape relationships.

Table E1: Firm Innovation and the Probability of Being Acquired

Dependent var:	Probability of being acquired (annual)			
Innovation var:	Patent count	Citations	Originality	Breakthrough
Innov	0.0025	0.0051**	0.0059***	0.0193**
	(1.55)	(2.62)	(2.71)	(2.44)
Innov^2	-0.0011***	-0.0007***	-0.0008***	-0.0018***
	(-4.83)	(-3.49)	(-4.68)	(-3.13)
$ln(Assets)$	0.0028***	0.0018***	0.0024***	0.0003
	(5.32)	(4.38)	(4.71)	(0.53)
Leverage	-0.0065**	-0.0043	-0.0057	-0.0004
	(-2.03)	(-1.31)	(-1.52)	(-0.08)
MB	-0.0001	-0.0001	-0.0001	-0.0002
	(-1.23)	(-1.01)	(-1.47)	(-1.15)
ROA	0.0027***	0.0033**	0.0034***	0.0050**
	(2.99)	(2.51)	(3.53)	(2.09)
Cash	0.0030	0.0061	0.0015	0.0083
	(0.50)	(1.08)	(0.33)	(1.00)
R&D	0.0246***	0.0255***	0.0224***	0.0275***
	(6.31)	(4.00)	(6.07)	(3.05)
Tangibility	-0.0087**	-0.0070	-0.0085*	-0.0030
	(-2.10)	(-1.63)	(-1.79)	(-0.45)
Div	-0.0085***	-0.0082***	-0.0084***	-0.0094***
	(-3.73)	(-3.49)	(-3.45)	(-2.76)
Age	0.0004	0.0005	0.0004	0.0011***
	(1.61)	(1.68)	(1.27)	(2.94)
Age^2	-0.0000	-0.0000	-0.0000	-0.0000***
	(-1.41)	(-1.64)	(-0.99)	(-2.72)
Const	0.0186***	0.0135**	0.0116	-0.0323
	(4.41)	(2.03)	(1.54)	(-1.17)
Industry F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	44665	41411	37875	22801
adjusted R^2	0.007	0.006	0.007	0.008

Notes: This table reports the regression results from Equation 1. *Innov* is the firm's innovation, measured by its patent count (column 1), patent citations (column 2), originality (column 3), and breakthrough innovations (column 4). All innovation variables are constructed as the stock of past innovations in the last three years. Patent count is the number of patents applied for by a firm in a given year. Patent citations are total citations received by patents applied for by a firm in a given year, which are corrected for truncation and technology class biases following Hall, Jaffe, and Trajtenberg (2001). Originality is the dispersion of technology classes cited by the firm's patents as described in Hall, Jaffe, and Trajtenberg (2001). Breakthrough innovations are defined as the number of patents that are in the top 10% of all patents according to the number of citations received among all patents applied for in that year as in Acemoglu, Akcigit, and Celik (2018). Other variable definitions are in Table 1. Standard errors are clustered by the firm's industry (Fama-French 48-industry classification).

TABLE E2: FIRM INNOVATION AND THE METHOD OF PAYMENT

Dependent var:	Equity share in acquisition offer			
Innovation var:	Patent count	Citations	Originality	Breakthrough
$Innov_{tar}$	0.0272	0.0290**	0.0356***	0.0175
	(1.43)	(2.07)	(2.77)	(0.49)
RRV_{acq}	0.1036***	0.1087***	0.0843*	0.1062**
	(2.94)	(2.89)	(2.03)	(2.07)
$Innov_{acq}$	-0.0035	-0.0014	-0.0152**	0.0193**
	(-0.26)	(-0.14)	(-2.04)	(2.28)
$Size_{acq}$	-0.0558***	-0.0592***	-0.0367***	-0.0809***
	(-3.23)	(-3.48)	(-2.77)	(-3.06)
MB_{acq}	0.0095	0.0094	0.0046	-0.0023
	(0.86)	(0.87)	(0.42)	(-0.12)
$Leverage_{acq}$	0.0164	-0.0024	-0.0338	-0.0744
	(0.18)	(-0.02)	(-0.38)	(-0.54)
ROA_{acq}	-0.3152*	-0.4045**	-0.3943**	-0.5848**
	(-1.73)	(-2.37)	(-2.07)	(-2.10)
RRV_{tar}	0.1010***	0.0960***	0.0904***	0.1281***
	(4.07)	(4.16)	(4.13)	(4.30)
$RelSize_{tar}$	0.0136	0.0080	-0.0010	0.0135
	(1.21)	(0.71)	(-0.08)	(0.45)
MB_{tar}	0.0027	0.0012	0.0108	-0.0154
	(0.25)	(0.10)	(1.19)	(-0.79)
$Leverage_{tar}$	0.0886	0.0722	0.0756	0.0093
	(0.82)	(0.72)	(0.66)	(0.07)
ROA_{tar}	-0.0068	0.0076	-0.0089	0.1186
	(-0.08)	(0.10)	(-0.13)	(1.32)
$Diversification$	-0.0145	-0.0175	0.0005	-0.0527
	(-0.28)	(-0.35)	(0.01)	(-0.90)
$Const$	1.3272***	1.3224***	1.0116***	1.5955**
	(3.97)	(4.12)	(4.05)	(2.58)
Industry F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	526	484	447	283
adjusted R^2	0.228	0.222	0.222	0.247

Notes: This table reports the regression results from Equation 2. *Innov* is the firm's innovation, measured by its patent count (column 1), patent citations (column 2), originality (column 3), and breakthrough innovations (column 4). All innovation variables are constructed as the stock of past innovations in the last three years. Patent count is the number of patents applied for by a firm in a given year. Patent citations are total citations received by patents applied for by a firm in a given year, which are corrected for truncation and technology class biases following Hall, Jaffe, and Trajtenberg (2001). Originality is the dispersion of technology classes cited by the firm's patents as described in Hall, Jaffe, and Trajtenberg (2001). Breakthrough innovations are defined as the number of patents that are in the top 10% of all patents according to the number of citations received among all patents applied for in that year as in Acemoglu, Akcigit, and Celik (2018). Other variable definitions are in Table 1. Standard errors are clustered by the acquirer's industry (Fama-French 48-industry classification).

TABLE E3: FIRM INNOVATION AND THE PROBABILITY OF DEAL COMPLETION

Dependent var:	Offer acceptance indicator			
Innovation var:	Patent count	Citations	Originality	Breakthrough
$Innov_{tar}$	-0.0365***	-0.0286***	-0.0292**	-0.0616***
	(-2.91)	(-2.83)	(-2.54)	(-3.03)
RRV_{acq}	-0.0257	-0.0274	-0.0079	-0.0235
	(-0.96)	(-1.00)	(-0.26)	(-0.59)
$Innov_{acq}$	-0.0009	0.0041	0.0008	-0.0012
	(-0.09)	(0.58)	(0.12)	(-0.14)
$Size_{acq}$	0.0363**	0.0331**	0.0360***	0.0397**
	(2.56)	(2.43)	(2.65)	(2.02)
MB_{acq}	0.0003	0.0028	0.0037	-0.0051
	(0.03)	(0.26)	(0.35)	(-0.30)
$Leverage_{acq}$	-0.0796	-0.0590	-0.0197	-0.1001
	(-0.90)	(-0.63)	(-0.20)	(-0.71)
ROA_{acq}	-0.2169	-0.2177	-0.1647	-0.0529
	(-1.54)	(-1.51)	(-1.12)	(-0.20)
RRV_{tar}	0.0436**	0.0476**	0.0266	0.0681**
	(2.09)	(2.19)	(1.19)	(2.44)
$RelSize_{tar}$	-0.0586***	-0.0578***	-0.0598***	-0.0682*
	(-3.63)	(-3.51)	(-3.16)	(-1.78)
MB_{tar}	-0.0214**	-0.0248**	-0.0119	-0.0251**
	(-2.32)	(-2.50)	(-1.19)	(-2.01)
$Leverage_{tar}$	-0.0526	-0.0701	-0.0673	0.0111
	(-0.79)	(-1.02)	(-0.94)	(0.12)
ROA_{tar}	-0.1372*	-0.1568*	-0.1812**	-0.1125
	(-1.71)	(-1.88)	(-2.08)	(-1.04)
$Diversification$	-0.0297	-0.0346	-0.0598*	0.0176
	(-0.91)	(-1.01)	(-1.67)	(0.38)
$Const$	0.5177	0.5908*	0.4444	0.7397
	(1.44)	(1.67)	(1.03)	(1.60)
Industry F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	589	547	490	322
adjusted R^2	0.119	0.115	0.127	0.125

Notes: This table reports the regression results from Equation 3. *Innov* is the firm's innovation, measured by its patent count (column 1), patent citations (column 2), originality (column 3), and breakthrough innovations (column 4). All innovation variables are constructed as the stock of past innovations in the last three years. Patent count is the number of patents applied for by a firm in a given year. Patent citations are total citations received by patents applied for by a firm in a given year, which are corrected for truncation and technology class biases following Hall, Jaffe, and Trajtenberg (2001). Originality is the dispersion of technology classes cited by the firm's patents as described in Hall, Jaffe, and Trajtenberg (2001). Breakthrough innovations are defined as the number of patents that are in the top 10% of all patents according to the number of citations received among all patents applied for in that year as in Acemoglu, Akcigit, and Celik (2018). Other variable definitions are in Table 1. Standard errors are clustered by the acquirer's industry (Fama-French 48-industry classification).

Table E4: Firm Innovation and the Probability of Being Acquired – (Inverted-U Hypothesis Test)

	Patent count	Citations	Originality	Breakthrough
lower bound				
t-value	2.390	3.573	2.906	4.495
P>|t|	0.011	0.000	0.003	0.000
upper bound				
t-value	-5.588	-4.548	-5.191	-5.674
P>|t|	0.000	0.000	0.000	0.000

Notes: To further check the robustness of the inverted-U relationship between the probability of being acquired and firm innovation stock, we test whether or not the slope of the fitted curve is positive at the start and negative at the end of the interval of firm innovation stock following Lind and Mehlum (2010). This table reports the hypothesis testing results.

TABLE E5: ELIMINATING INFORMATION FRICTIONS IN THE EXTENDED MODEL WITH TARGET-SIDE LACK OF INFORMATION ($\Upsilon = 4.57\%$)

	Extended Model	Short-Term Effect ($\omega = 0$) baseline $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$	Long-Term Effect ($\omega = 0$) new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$
market value	11.056	11.559	11.345
standalone value	10.387	10.387	10.305
capitalized expected gain from M&A	0.669	1.172	1.040
capitalized expected gain from M&A/market value	6.05%	10.14%	9.17%
avg. R&D intensity	5.85%	5.85%	6.42%
aggregate output	4.462	4.462	4.576
avg. merger probability	1.67%	2.28%	1.95%
consumption	2.010	2.010	2.062
avg. firm growth rate	8.64%	8.44%	8.87%

Notes: This table reports model implications in the extended model with target-side lack of information in which we set the parameter $\Upsilon = 0.5 * 9.146\% \approx 4.57\%$ while keeping other parameters the same as the baseline estimates, and in a counterfactual economy in which information frictions are eliminated by setting $\omega = 0$. Column one shows the baseline results. In column two, we evaluate the short-term effect of information frictions, keeping firm innovation policy and the cross-sectional distribution as they are in the baseline equilibrium. Column three reports the long-term effect in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Market value is the model-implied value of the firm. We decompose the market value into two components: a standalone value derived by shutting down a firm's opportunity to participate in the M&A market during its lifetime, and an option value of gaining from the M&A market. R&D intensity is measured as the total expenses firms invest on innovation divided by firm assets; output is the aggregate production in the economy; avg. merger probability is the total number of mergers divided by the total number of firms in the economy; firm growth rate is the average firm sale growth.

Table E6: Eliminating Information Frictions in the Extended Model with Target-side Lack of Information ($\Upsilon = 9.15\%$)

	Extended Model	Short-Term Effect ($\omega = 0$) baseline $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$	Long-Term Effect ($\omega = 0$) new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$
market value	11.064	11.542	11.342
standalone value	10.382	10.382	10.306
capitalized expected gain from M&A	0.681	1.159	1.035
capitalized expected gain from M&A/market value	6.16%	10.04%	9.13%
avg. R&D intensity	5.87%	5.87%	6.42%
aggregate output	4.464	4.464	4.574
avg. merger probability	1.68%	2.25%	1.94%
consumption	2.011	2.011	2.060
avg. firm growth rate	8.65%	8.47%	8.87%

Notes: This table reports model implications in the extended model with target-side lack of information in which we set the parameter $\Upsilon = 9.15\%$ while keeping other parameters the same as the baseline estimates, and in a counterfactual economy in which information frictions are eliminated by setting $\omega = 0$. Column one shows the baseline results. In column two, we evaluate the short-term effect of information frictions, keeping firm innovation policy and the cross-sectional distribution as they are in the baseline equilibrium. Column three reports the long-term effect in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Market value is the model-implied value of the firm. We decompose the market value into two components: a standalone value derived by shutting down a firm's opportunity to participate in the M&A market during its lifetime, and an option value of gaining from the M&A market. R&D intensity is measured as the total expenses firms invest on innovation divided by firm assets; output is the aggregate production in the economy; avg. merger probability is the total number of mergers divided by the total number of firms in the economy; firm growth rate is the average firm sale growth.

Table E7: Eliminating Information Frictions in the Extended Model with Target-side Lack of Information ($\Upsilon = 18.29\%$)

	Extended Model	Short-Term Effect ($\omega = 0$) baseline $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$	Long-Term Effect ($\omega = 0$) new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$
market value	11.079	11.507	11.335
standalone value	10.374	10.374	10.309
capitalized expected gain from M&A	0.706	1.133	1.026
capitalized expected gain from M&A/market value	6.37%	9.85%	9.05%
avg. R&D intensity	5.89%	5.89%	6.41%
aggregate output	4.469	4.469	4.569
avg. merger probability	1.70%	2.21%	1.94%
consumption	2.013	2.013	2.058
avg. firm growth rate	8.68%	8.52%	8.87%

Notes: This table reports model implications in the extended model with target-side lack of information in which we set the parameter $\Upsilon = 2 * 9.146\% \approx 18.29\%$ while keeping other parameters the same as the baseline estimates, and in a counterfactual economy in which information frictions are eliminated by setting $\omega = 0$. Column one shows the baseline results. In column two, we evaluate the short-term effect of information frictions, keeping firm innovation policy and the cross-sectional distribution as they are in the baseline equilibrium. Column three reports the long-term effect in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Market value is the model-implied value of the firm. We decompose the market value into two components: a standalone value derived by shutting down a firm's opportunity to participate in the M&A market during its lifetime, and an option value of gaining from the M&A market. R&D intensity is measured as the total expenses firms invest on innovation divided by firm assets; output is the aggregate production in the economy; avg. merger probability is the total number of mergers divided by the total number of firms in the economy; firm growth rate is the average firm sale growth.

TABLE E8: ELIMINATING INFORMATION FRICTIONS IN THE EXTENDED MODEL WITH TARGET-SIDE LACK OF INFORMATION ($\Upsilon = 27.44\%$)

	Extended Model	Short-Term Effect ($\omega = 0$) baseline $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$	Long-Term Effect ($\omega = 0$) new $F_s(z)$ and $\hat{\mu}_i(z;\Theta)$
market value	11.100	11.480	11.328
standalone value	10.351	10.351	10.311
capitalized expected gain from M&A	0.749	1.129	1.017
capitalized expected gain from M&A market value	6.75%	9.84%	8.98%
avg. R&D intensity	5.93%	5.93%	6.40%
aggregate output	4.484	4.484	4.564
avg. merger probability	1.73%	2.17%	1.93%
consumption	2.019	2.019	2.056
avg. firm growth rate	8.71%	8.57%	8.87%

Notes: This table reports model implications in the extended model with target-side lack of information in which we set the parameter $\Upsilon = 3 * 9.146\% \approx 27.44\%$ while keeping other parameters the same as the baseline estimates, and in a counterfactual economy in which information frictions are eliminated by setting $\omega = 0$. Column one shows the baseline results. In column two, we evaluate the short-term effect of information frictions, keeping firm innovation policy and the cross-sectional distribution as they are in the baseline equilibrium. Column three reports the long-term effect in which we allow firm distribution to evolve endogenously and let firms reoptimize their innovation policies. Market value is the model-implied value of the firm. We decompose the market value into two components: a standalone value derived by shutting down a firm's opportunity to participate in the M&A market during its lifetime, and an option value of gaining from the M&A market. R&D intensity is measured as the total expenses firms invest on innovation divided by firm assets; output is the aggregate production in the economy; avg. merger probability is the total number of mergers divided by the total number of firms in the economy; firm growth rate is the average firm sale growth.

Table E9: Sensitivity of Moments to Parameters

	ω	Λ	χ	α	$\underline{\mu}$	σ	ρ	ψ	γ
loading of equity ratio on innov stock	0.336	-2.580	0.001	-1.123	-3.006	-0.077	-0.589	-2.099	-0.219
average equity ratio	0.830	-1.097	0.000	0.618	-0.185	7.210	4.979	2.042	0.474
loading of deal completion on innov stock	-0.066	0.202	-0.001	-0.358	-0.494	0.058	0.184	-1.944	-0.021
average value loss	0.013	0.093	0.000	0.002	-0.013	0.117	-0.009	0.079	0.008
average R&D intensity	-0.005	0.000	-0.002	0.445	2.478	0.477	1.102	-1.668	0.091
relative size	0.095	0.000	0.000	-0.338	-0.431	-2.894	-0.150	-0.908	-0.050
firm growth rate	-0.005	0.000	0.000	0.322	0.219	0.325	0.227	0.002	0.057
coefficient of variation of $ln(ME)$	-0.014	-0.002	0.000	0.721	-0.188	0.594	1.595	1.361	0.122
autocorr. of $ln(ME)$	-0.010	-0.004	0.000	0.438	-0.112	-0.089	0.629	-1.270	0.075
firm entry rate	-0.005	0.000	0.000	0.284	0.119	-0.074	0.201	0.296	0.050
average merger probability	-0.005	0.000	0.000	0.284	0.119	-0.074	0.201	-0.704	0.050
average realized gain	-0.018	-0.109	0.000	0.209	-0.399	0.071	0.190	0.805	0.016
loading of merger prob. on innov stock	-0.051	0.178	0.000	0.157	0.126	-0.015	0.098	0.115	0.037
loading of merger prob. on innov stock^2	0.004	-0.031	0.000	-0.011	-0.011	0.006	-0.004	-0.004	-0.002

Notes: This table shows the sensitivity of model-implied moments (in rows) with respect to model parameters (in columns). Moments are defined in detail in Section 5.1.