The Aggregate Implications of Mergers and Acquisitions*

Joel M. David†

University of Southern California

August 30, 2017

Abstract

This paper develops a search and matching model of mergers and acquisitions (M&A) and uses it to evaluate the implications of merger activity for aggregate economic outcomes. The theory is consistent with a rich set of micro-level facts on US M&A, including, e.g., sorting among merging firms, a substantial merger premium and serial acquisition. It provides a sharp link between these facts and the nature of merger gains. Estimating the model shows that both synergies, i.e., complementarities between merging firms, and productivity improvements of target firms are important in generating firm-level gains. Gains are split almost equally between the parties, despite large differences in abnormal stock returns. The estimated model suggests that M&A has a significant beneficial impact on aggregate outcomes - the results show a contribution to steady state output of 13% and 4% for consumption. Endogenous sorting plays an important role in generating these gains, which occur through the reallocation of resources across firms and changes in firm selection and new entrepreneurship. The economy is not efficient, suggesting a scope for policy improvements.

Keywords: Mergers & Acquisitions, Search & Matching, Reallocation, Aggregate Productivity

*I am grateful to the editor, Philipp Kircher, and three anonymous referees for their valuable comments and suggestions, and to Hugo Hopenhayn for his invaluable guidance. I also thank Andy Atkeson, Pierre-Olivier Weill, Andrea Eisfeldt and many seminar and conference participants for their insightful comments.

†Email: joeldavi@usc.edu.

1 Introduction

A large and growing body of work demonstrates the importance of resource reallocation across firms in determining aggregate economic outcomes. Mergers and acquisitions represent an important vehicle for such reallocation in the US economy and account for huge flows of resources between firms.[1] A recent literature uses stock market data to document clear empirical patterns at the micro-level in US M&A and uses these facts to make inferences regarding the possible motives for merger. In important contributions, Jovanovic and Rousseau (2002) point out a high-buys-low pattern and propose a "q-theory of mergers" in which mergers serve to transfer resources from low to high productivity firms, whereas Rhodes-Kropf and Robinson (2008) document a like-buys-like pattern, suggesting synergistic forces. In this paper, I build on these insights to investigate the motives and gains from merger, both for the transacting firms and the aggregate economy. I develop a quantitative model of M&A that is at once consistent with the rich set of micro-level facts and tractable enough to make explicit the mechanisms through which firm outcomes from M&A aggregate to impact the macroeconomy through the reallocation of resources across firms and changes in firm selection and new entrepreneurship.

The starting point of my analysis is to establish a set of facts on merging firms using a broad set of measures from firm-level operating data (e.g., sales, employment, profitability, etc.), in addition to the stock market data previously examined in the literature. The key findings are: (1) acquiring firms are generally larger and more profitable than their targets; (2) there is a large degree of positive assortative matching between transacting firms; and (3) acquirers tend to be the largest and most profitable firms, but targets are not the smallest or least profitable. Additionally, I document the prevalence of repeat acquisition and show that data on the pricing of transactions, i.e., merger premia, are informative about merger gains and split. Given the robustness of these patterns across a wide set of measures of firm fundamentals (in addition to stock market valuations), one conclusion that can be drawn is that any quantitative model of M&A must be able to jointly reconcile them.

The second contribution of the paper is to develop a theory that matches these facts. My point of departure is a parsimonious model of M&A through search, matching and bargaining in the spirit of Rhodes-Kropf and Robinson (2008), which builds on the underlying structure in Shimer and Smith (2000).[2] Heterogeneous firms engage in costly and potentially time-

[1] For example, expenditures on M&A from 1980-2009 averaged about 5% of GDP annually, reaching as high as 16% in the late 1990s, and about 44% of de novo business investment. The rate of capital reallocation via M&A has also averaged about 5%, accounting for an annual average of about two-thirds of total capital reallocation among large US firms, a figure that has grown to over 80% (Data are from SDC Platinum, Bureau of Economic Analysis, and for total reallocation, downloaded from Andrea Eisfeldt's website).

[2] Martos-Vila (2008) is another example of recent work recognizing the importance of search frictions in M&A. In describing the merger market, the corporate finance literature generally highlights the importance of

consuming search for profitable merger partners. Upon meeting, the parties bargain over the gains and decide whether to consummate the transaction or not. The set of profitable partners is determined by a "merger technology" determining the characteristics, and hence profitability, of the continuing entity as a function of those of the two pre-merger firms. I build on this framework in a number of key dimensions. First, I generalize the technology to a flexible Cobb-Douglas form. This specification can capture both synergistic and q-theory motives for merger and is able to match the rich set of micro-level patterns. In contrast, each of these theories is consistent with some aspects of observed merger activity, but on its own cannot explain the full set of empirical facts. A second novel ingredient is the addition of repeat acquisition. This is an important feature of the merger market, where many firms make multiple acquisitions (and one-time acquirers often themselves become targets), and, as I discuss below, has important implications for the sorting patterns implied by the theory. Thus, both of these elements play crucial roles in reconciling merger gains with observed firm matching.

The third main contribution of the paper is to analyze the interactions of M&A with the aggregate economy, specifically, the effects on entrepreneurship, selection and the size distribution of firms. I embed the theory of firm-level merger activity into a general equilibrium model of industry dynamics in the spirit of Hopenhayn (1992) and Melitz (2003). The framework yields simple expressions for macroeconomic variables, e.g., aggregate productivity, which is in part endogenous and depends on merger market outcomes.[3] M&A influences aggregate outcomes, first, by reshaping the distribution of firm productivity and allocation of resources across those firms and second, by changing the incentives for firm entry and exit, i.e., through the extent of new entrepreneurship and selection. To the best of my knowledge, my analysis is the first to shed light on the aggregate implications of M&A in a standard macroeconomic framework.

The presence of repeat acquisition is a key departure from the standard search paradigm. In particular, given that firms are able to rematch and so usual option value arguments do not apply, conditions for sorting to arise in equilibrium are not clear. Indeed, simulations show that with rematching, standard assumptions on complementarities are no longer sufficient. I use a two-period model in the spirit of Eeckhout and Kircher (2011) to derive a set of necessary and sufficient conditions for sorting in this environment - the technology must exhibit (i) a sufficient degree of complementarities or "synergies," and (ii) sufficient curvature with respect to each firm's individual contribution. In the Cobb-Douglas specification I work with, these conditions correspond to (i) an overall degree of curvature exceeding one and (ii) an elasticity with respect to each argument less than one.[4] The curvature condition is novel to this setting and stems

search, matching and bargaining, for example, DePamphilis (2009) and Welch (2009).

[3]In contrast, productivity growth is exogenous in those papers. For example, in Melitz (2003), firm productivity is constant from birth until death.

[4]Condition (i) is also a statement about the magnitude of the cross-partial.

from the presence of repeat acquisition. The lower bound on complementarities comes from the fact that firms have heterogeneous standalone profits that, for a poor match, can exceed post-merger profits.[5] Numerical simulations suggest these conditions remain necessary in the infinite-horizon case.

I estimate the model by matching key moments of observed M&A activity, e.g., the characteristics of merging firms and the merger premium. The estimated model is able to replicate the broad set of empirical facts. At the micro-level, the parameter estimates imply the potential for sizable gains to transacting firms - on average, about 7% of combined value - with important roles for both merger synergies and productivity improvements on transferred resources along the lines of the q-theory. Strikingly, the results imply an almost equal split of the gains, despite the fact that the model predicts large disparities in abnormal stock returns (in favor of the target), as observed in the data.

I use the estimated model to evaluate the impact of M&A on the aggregate economy. First, I compare the aggregate outcomes in the estimated economy to a counterfactual one with no merger activity. The results show a significant beneficial impact of M&A, contributing about 13% to steady state output and 4% to consumption. Importantly, I find that a significant portion of the gains arises from general equilibrium channels that would be difficult to measure using only data on observed mergers, specifically, changes in the extent and selection of new entrepreneurs that enter the economy. The endogenous sorting of firms plays an important role in generating aggregate gains - for example, firms' endogenous search decisions account for almost half of the gains from M&A and completely random matching would lead to aggregate outcomes that are worse than having no merger activity at all. While the presence of M&A leads to increases in aggregate productivity, the results reveal an important tradeoff for households - much of the productivity gains are offset by an increase in costly churn among entrants, i.e., more unsuccessful attempted entry by new entrepreneurs, each of whom uses resources to pay entry costs that otherwise would have gone to households.

To analyze efficiency of the equilibrium, I characterize the solution to the problem of a social planner constrained by the same search frictions. Along the lines of Shimer and Smith (2001), I prove that in the presence of heterogeneity, the economy is not efficient. The planner's optimality conditions reveal both a "thick markets" externality - firms make search decisions taking into account only their own share of match surplus, whereas the planner accounts for the total surplus, and a "congestion" externality - firms do not internalize that their own search decisions make it more difficult for others to match. I illustrate these effects using a number of simple policies.

The paper is organized as follows. Section 2 presents the key empirical facts. Section 3 devel-

[5] In many search models, the profits of unmatched agents are constant, or set to zero.

ops the model and analytic results. Section 4 describes the estimation and Section 5 evaluates the aggregate implications. Section 6 concludes and discusses directions for future research. Details of data work, proofs and a number of extensions/robustness checks are provided in the Appendix and Online Appendix.

Related literature. This paper relates to several branches of literature. First, I build closely on the seminal contributions of Jovanovic and Rousseau (2002) and Rhodes-Kropf and Robinson (2008) already discussed. These papers use stock market data to identify high-buys-low and like-buys-like patterns in M&A, respectively, and use these facts to shed light on the nature of merger gains.[6] In important departures, I introduce a general specification of the merger technology and allow for repeat acquisition. With these elements, the model is able to fit the full set of empirical facts and I can make quantitative inferences regarding the parameters of the technology and the associated gains from merger. Indeed, one contribution of my work is to provide a framework that jointly captures the forces for merger uncovered in these papers. I use these results to provide new insights into the effects of M&A on the entry decisions of entrepreneurs, selection and the size distribution of firms, features that are new to my setting.

There is a large body of work focusing on measuring the gains and split from merger using event-study analysis of financial market data and a smaller number of papers that investigate gains by examining the pre- and post-merger performance of the transacting firms. Comprehensive reviews of these studies are in Andrade et al. (2001), Betton et al. (2008) and Eckbo (2014). Much of the recent work along these lines find potentially significant returns to firms from merging using both stock market data and direct measures of operating performance. Section 4.3 discusses the findings of this literature in more detail. Compared to this approach, the structural model I develop allows me to more fully investigate the economic forces creating merger gains and explore the interaction of M&A with aggregate outcomes. The results suggest that there can be important indirect effects of M&A in general equilibrium that would be challenging to measure using only data on observed transactions.

A number of recent papers build on Shimer and Smith (2000) to derive relationships between the match production function and the resulting distribution of matches in environments with search frictions, important examples of which include Atakan (2006), Eeckhout and Kircher (2010) and Lentz (2010). Eeckhout and Kircher (2011) and Hagedorn et al. (2017) explore the identification of sorting in environments where agent types are not observable. An excellent recent review of this line of work is in Chade et al. (2017). My analysis builds on the key insights of this literature to link observed sorting patterns among merging firms to the nature

[6] Papers relying on versions of the q-theory of Jovanovic and Rousseau (2002) to study reallocation include Eisfeldt and Rampini (2006), Eisfeldt and Rampini (2008), Jovanovic and Rousseau (2008), and Faria (2008).

of merger gains. I show that the conditions for sorting may differ in environments with repeat matching - specifically, assumptions on complementarities may not be sufficient; the curvature of the match production function plays an important role as well.

Lastly, my theory of M&A as one way in which knowledge is transmitted across firms relates to a recent literature on knowledge diffusion through imitation and long-run growth, e.g., Lucas and Moll (2014) and Perla and Tonetti (2014). Compared to those papers, I study a specific setting for the transfer of knowledge in which both agents search for profitable matching opportunities, bargain over the surplus and combine using a general technology, which can be empirically disciplined using micro-level facts on the sorting patterns of matching agents.[7] This link suggests that the approach here may be relevant for that literature - specifically, the fact that particular theories of how knowledge spreads can have strong implications for exactly how agents match and sort may be useful to guide future empirical work along those lines.

2 Empirical Patterns in US M&A

In this section, I combine transaction and firm-level operating data to establish a number of facts on US merger activity. The key empirical findings are: (1) acquiring firms are generally larger and more profitable than their targets; (2) there is a large degree of positive assortative matching between transacting firms; (3) acquirers tend to be the largest and most profitable firms, but targets are not the smallest or least profitable; (4) repeat acquisition is prevalent; and (5) most mergers are small and merger premia are substantial.

2.1 The Data

I develop a sample of almost 58,000 domestic transactions announced between 1977 and 2009. Transaction-level data are from the Thomson Reuters SDC Platinum database (SDC). SDC is a comprehensive source of data on US M&A, covering all corporate transactions involving at least 5% of the ownership of a company where the transaction is valued at $1 million or more (after 1992, all deals are covered) or where the value of the transaction was undisclosed. SDC covers both public and private transactions.[8] Deal characteristics contained in SDC include the transaction value (purchase price) and premium, which is defined as the percentage by which the purchase price exceeds the pre-merger market value of the target. SDC contains a

[7] Indeed, Section 3.2 proves that the setup in Lucas and Moll (2014) is a variant of my framework where the technology is purely imitative, search is one-sided on the part of targets and targets capture the entirety of the surplus from a match.

[8] I restrict the sample to transactions valued over $1 million. Datasets and sample construction are described in detail in Appendix A.

number of pre-transaction statistics on the merging firms including sales, employment, property, plant, and equipment (PP&E), earnings before interest, taxes, depreciation, and amortization (EBITDA), and market value, which are generally calculated for the 12 month period preceding the deal announcement.[9] PP&E and EBITDA capture the size of the firm's capital stock and level of profitability, respectively. I deflate all nominal variables to constant 2005 dollars using the CPI.

To compare the characteristics of transacting firms to the population of firms, I obtain the corresponding set of statistics for the universe of Compustat firms. To ensure comparability between the two sets of firms, I match the SDC database to Compustat and use Compustat operating statistics in calculations that involve industry aggregates, e.g., industry means or medians. I use the SDC statistics in calculations that do not, mainly because SDC provides wider coverage of private firms, greatly expanding the set of included firms, in particular targets. For example, sales are available for about 6,800 targets from Compustat and for 18,500 targets from SDC. This also increases the number of observations compared to using data on stock market valuations, which are available for about 6,900 targets.

2.2 Transaction Sizes and Prices

Most deals are small, premia are large. Table 1 reports summary statistics of transaction values and premia. The merger premium is defined as the percentage by which the purchase price exceeds the pre-merger market value of the target firm. The mean transaction value is quite modest at $267 million as is the median at only $31 million. The difference reflects a great deal of right-skewness in the distribution of transaction values, i.e., the majority of mergers are small, with some very large outliers. In contrast, merger premia tend to be substantial, with a mean premium of about 47% and a median of 38%.[10] The typical transaction is characterized by a purchase price well above the pre-merger market value of the target firm - on average, the acquirer pays almost 50% more than the standalone value of the firm it is purchasing.

2.3 Joint and Marginal Distributions of Transacting Firms

Figure 1 plots the entire joint distribution over the profitability (i.e., EBITDA) of acquirers and targets. Each point in the figure represents one transaction. In Panel A, I show profits as

[9] Market values are calculated 4 weeks prior to announcement. Firm-level data are only available for a subset of transactions, in large part because many of the firms in the database are privately owned and are not required to report operating statistics to any regulatory agency.

[10] To avoid the known run-up in target share price once rumors of the merger begin to circulate, the premium is calculated using the market value of the target firm 4 weeks prior to the merger announcement. Following the literature, for example, Officer (2003), Moeller et al. (2004) and Malmendier et al. (2016), I truncate premia below zero and above 200%.

Table 1: Transaction Values and Premia

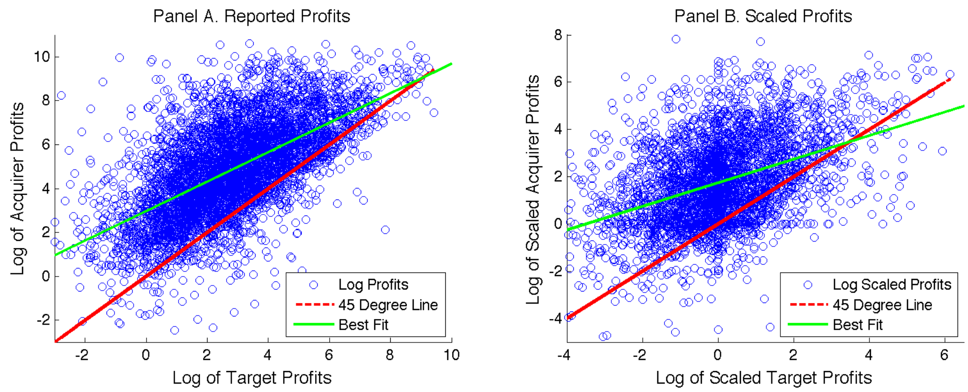
	Trans. Val. ($M)	Premium (%)
Mean	267.4	46.8
Median	31.0	38.2
SD	1,911.5	34.8
Max	186,824.1	200.0
Min	0.9	0.0
N	57,858	5,868

Notes: Table reports summary statistics of transaction values and premia in US M&A over the period 1977-2009. Transaction values are deflated to 2005 dollars using the CPI. Data are from SDC Platinum.

reported. In Panel B, I rescale the data by deviating the profits of each firm from the median in its industry. I make this adjustment to ensure that cross-industry differences between the transacting firms do not skew the results. Industries are defined at the 4-digit SIC level as reported in Compustat. There are about 5,100 transactions in the first panel and 3,500 in the second. The figure reveals two key facts, each of which is robust to rescaling: first, the high-buys-low pattern in stock market valuations documented in Andrade et al. (2001) and Jovanovic and Rousseau (2002) manifests itself in production-side variables. The figure shows significant and pervasive profitability differences between acquirers and targets - the majority of transactions lie above the 45° line, as does the line of best fit. Second, the like-buys-like pattern in Rhodes-Kropf and Robinson (2008) shows up as well - there is strong positive sorting among matching firms, a feature highlighted by the positive slope of the line of best fit, as well as by the absence of points in the northwest corner of the plot, i.e., the most profitable acquirers do not tend to match with the least profitable targets, and a similar absence in the southeast corner, i.e., the least profitable acquirers do not match with the most profitable targets.[11] This section shows that these findings hold across a wide array of production-side variables.

Acquirers are larger and more profitable than their targets. Table 2 reports the mean and median log differences between matched acquirers and targets in profits, sales, employment, capital stock, and market value. Following the convention discussed, I report statistics both as reported and after rescaling by industry medians. The first set of columns shows that across all measures, acquirers are generally larger and more profitable than their targets. The disparities are substantial, with a mean and median difference between acquirer and target of about 2 log points, a factor of almost 7.5. Moreover, the differences are pervasive, with the acquirer

[11] A similar figure obtains when using other measures of firm size or profitability. For example, Online Appendix A shows the analogous figure for firm sales, for which the most observations are available. Broadly similar patterns obtain using measures of labor productivity and profit margins.



Notes: Panel A displays the log of acquirer and target profits in individual transactions. Panel B displays the log of profits after rescaling by deviating each firm from the median in its industry, i.e., log (acquirer profits) − log (median profits in acquirer's industry) for acquirers and log (target profits) − log (median profits in target's industry) for targets. Data are from SDC Platinum and Compustat.

Figure 1: Joint Distribution of Acquirers and Targets

exceeding its target in size and profitability in about 90% of transactions. The second set of columns confirm that these facts are robust to rescaling.

Table 2: Log Differences in Matched Acquirers and Targets

	Reported			Scaled by Industry Medians		
	Mean	Median	%>0	Mean	Median	%>0
Profits	2.1	1.9	90.0	1.7	1.6	83.2
Sales	2.0	1.9	89.0	1.6	1.4	81.7
Employment	2.0	1.8	87.1	1.4	1.3	79.4
Capital Stock	2.1	1.9	88.2	1.6	1.5	79.9
Market Value	2.3	2.1	95.3	1.9	1.7	86.4

Notes: The left-hand panel reports mean and median log differences in the characteristics of matched acquirers and targets across individual transactions, e.g., log (acquirer profits) − log (target profits), and the percent of transactions in which the acquirer exceeds the target. The right-hand panel reports mean and median log differences after rescaling by deviating each firm from the median in its industry, e.g., [log (acquirer profits) − log (median profits in acquirer's industry)] − [log (target profits) − log (median profits in target's industry)], and the percent of transactions in which the acquirer exceeds the target after rescaling. Data are from SDC Platinum and Compustat.

Positive assortative matching. The second fact revealed in Figure 1 is the pattern of positive sorting between acquirers and targets - more profitable acquirers tend to partner with more profitable targets, as in turn do the less profitable. Table 3 reports correlations between the characteristics of acquirers and targets in individual matches across metrics of firm size and profitability. Again, I report the correlations from the data as reported and after rescaling. The first column shows a large positive correlation between acquirers and targets along all dimensions of the data, on the order of about 0.6 or higher, confirming the strength and ubiquity of positive sorting. The second column confirms this pattern is robust to rescaling.

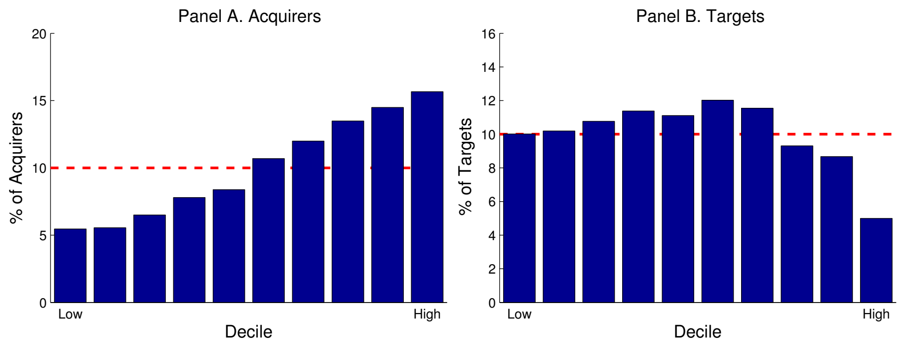
Table 3: Log Correlations Between Acquirers and Targets

	Reported	Scaled
Profits	0.63	0.43
Sales	0.62	0.42
Employment	0.58	0.38
Capital Stock	0.69	0.39
Market Value	0.64	0.48

Notes: Table reports log correlations of acquirer and target characteristics. Scaled variables are deviations from industry medians, i.e., log (acquirer profits) − log (median profits in acquirer's industry) for acquirers and log (target profits) − log (median profits in target's industry) for targets. Data are from SDC Platinum and Compustat.

Acquirers are large and profitable, targets are median. Figure 2 plots the marginal distributions of acquirer and target firms. Specifically, I calculate the deciles of the firm size distribution measured by profitability across all firms in each industry and year. I count the proportion of acquirers and targets that fall into each decile. If transacting firms were distributed similarly to the population of firms (in their industries), about 10% of transacting firms would fall into each decile, which is the dashed line. Deciles above this line are overrepresented, i.e., they account for more than 10% of transacting firms, and deciles below the line are underrepresented. The figure reveals that, in line with the findings on stock market valuations in Rhodes-Kropf et al. (2005) and Rhodes-Kropf and Robinson (2008), acquirers tend to be the most profitable firms but targets are not the least profitable. Panel A shows that acquirers predominantly come from the upper deciles of the distribution. Only about 5% of acquirers come from the lowest decile, with the proportion monotonically increasing to about 16% in the highest. The bottom 5 deciles are all underrepresented among acquiring firms while the top 5 are all overrepresented. Panel B reveals quite a different pattern for targets. Targets predominantly come from the middle of the distribution, deciles 3 to 7, are just about proportionally represented in deciles 1-2 and underrepresented in deciles 8-10, severely so in the last. Targets are not overrepresented in the lowest deciles, and indeed, are distributed approximately equally around the median firm so that the typical target is similar to the typical firm.

Table 4 confirms that these facts hold across a broader set of measures of firm performance. The table reports the mean and median log differences between acquirers and targets and the median firm in their respective industries across metrics of firm size and profitability. Acquirers



Notes: Figure displays the proportion of acquiring and target firms that fall into each decile of the firm size distribution (measured by profitability). Data are from SDC Platinum and Compustat.

Figure 2: Marginal Distributions of Transacting Firms

tend to be larger and more profitable than the median firm in their industry along every dimension. The mean and median differences are both substantial, hovering around 0.7 and 0.6 log points respectively, a factor of about 2, and the share of transactions with this feature is about two-thirds. In contrast, the typical target is almost identical to the median firm. Targets on average exceed the median firm in their industry on 3 out of 5 dimensions and are smaller on two. The absolute size of these differences are relatively small, especially as compared with the magnitudes by which acquirers differ from the median and by which acquirers exceed targets. On all dimensions, targets are distributed almost equally around the median firm so that the median difference between targets and the median firm is essentially zero.

Table 4: Log Deviations from Industry Median

	Acquirer			Target		
	Mean	Median	%>0	Mean	Median	%>0
Profits	0.74	0.57	64.1	-0.07	-0.01	45.7
Sales	0.75	0.58	64.5	0.14	0.00	49.4
Employment	0.57	0.50	63.1	0.10	0.00	48.8
Capital Stock	0.79	0.51	63.8	0.11	0.00	49.7
Market Value	1.01	0.86	69.0	-0.01	0.00	47.1

Notes: Table reports mean and median log differences in the characteristics of acquirers and targets and the median firm in their industries, e.g., log (acquirer profits) − log (median profits in acquirer's industry) for acquirers and log (target profits) − log (median profits in target's industry) for targets, and the percent of transactions in which acquirers/targets exceed the median. Data are from SDC Platinum and Compustat.

2.4 Repeat Acquisition

Repeat acquisition is an important feature of the merger market. Indeed, Table 5 shows that serial acquisition is not the exception, but rather the rule. The table reports the distribution of transactions by the number of purchases made by the acquirer during the sample period. In only about 19,000 transactions, or one-third the total, is the acquirer a one-time purchaser. In two-thirds of transactions, the acquirer is a repeat participant in the market. In about a quarter of transactions, the acquirer has made either 2 or 3 purchases, in another quarter between 4 and 10, and there is a long right tail, which I have truncated at 40, but which reaches to some acquirers that make over 100 purchases during the sample period.

Table 5: Distribution of Transactions by Number of Acquirer Purchases

Number of Purchases	Firms	Transactions	Share of Transactions
1	18,870	18,870	32.5
2-3	5,777	13,301	23.0
4-5	1,669	7,345	12.7
6-7	733	4,695	8.1
8-10	490	4,311	7.5
11-15	340	4,251	7.3
16-20	119	2,100	3.5
21-30	79	1,914	3.3
31-40	22	751	1.3
More than 40	6	320	0.5
Total	28,105	57,858	100.0

Notes: Table reports the distribution of transactions and firms by the number of purchases made by that firm, and the share of total transactions accounted for by each category. Data are from SDC Platinum and Compustat.

2.5 Measuring Merger Gains from Prices and Values

The merger premium reported in Table 1 and the value differences in Table 2 give a useful bound on the gains to firms from merging. Let V_a and V_t denote the pre-merger market values of the acquirer and target, respectively. By definition, the premium is equal to $prem_{a,t} = \frac{P_{a,t} - V_t}{V_t}$ where $prem_{a,t}$ denotes the premium when acquirer a purchases target t and $P_{a,t}$ the associated price. Without loss of generality, we can derive a formula for the combined percentage gains as

$$P_{a,t} = V_t + (1 - \beta_{a,t})\,\Sigma_{a,t} \quad \Rightarrow \quad \frac{\Sigma_{a,t}}{V_a + V_t} = \frac{prem_{a,t}}{1 - \beta_{a,t}} \frac{V_t}{V_a + V_t} \tag{1}$$

where $\Sigma_{a,t} = V_m - V_a - V_t$ denotes the combined dollar gains and $\beta_{a,t}$ the share that goes to the target. An assumption-free lower bound is when $\beta_{a,t} = 0 \; \forall \; a, t$. In this case, the target receives the entirety of the gains and the premium perfectly reflects the size of those gains. Using the mean value for the premium from Table 1 and the mean value differences from Table 2 gives a lower bound on mean gains of about 4.3% (the median is similar at 4.2%). Computing the right hand side of expression (1) for each transaction in the data and then averaging gives a somewhat higher value of 6.6%. These figures are lower bounds, since we have assumed that targets receive the totality of the gains. We can go further if we are willing to assume a bargaining structure, e.g., Nash. If firms split the gains equally, i.e., $\beta = \frac{1}{2}$ (which will be very close to the empirical results below), analogous calculations give average gains of between 8.5% and 13.1% (and median gains of 5.8%-8.3%).[12]

In sum, data on merger premia and pre-merger values suggest that there are significant gains to firms from merging, although a precise number requires having a value for β. However, although useful, these calculations cannot speak to (1) the economic forces that lead to these gains and so spur merger activity, and (2) the effect on the aggregate economy, since there may be indirect effects operating in general equilibrium, i.e., through changes in entry/exit and the size distribution of firms, that augment or mitigate these values.

3 The Model

I consider an infinite-horizon economy set in continuous time. There is a continuum of firms that are heterogeneous over some firm-specific intangible, z, which I will generally call productivity, but can alternatively be interpreted as organization capital, or the quality of the firm's management or product offering. The firm's z may evolve through merger activity and is otherwise persistent. Market competition results in variable profits that are directly proportional to z, such that higher productivity firms are more profitable, an assumption clarified in Section 3.3. For now I make no additional assumptions on the precise market structure. In order to produce and remain in operation, firms must pay a fixed cost of operation, c_f. Denote by Πz the variable profits of a firm of type z, where Π is a factor of proportionality common across firms. The firm's net profit flows are equal to

$$\pi(z) = \Pi z - c_f$$

[12]In principle, we could add data on post-merger values to jointly measure the gains and split, which is the event study strategy. I discuss this approach in Section 4.3.

3.1 The Merger Market

Firms participate in the merger market seeking value-increasing opportunities to combine with a merger partner. Through this process, firms are able to influence their z's, i.e., improve their productivity or quality, with an associated change in their level of profitability. Intuitively, the merger process enables firms to enhance their intangible assets by growing their organization capital or productive know-how, or bundling products in some quality improving way. Another interpretation would be of firms trading the blueprints or knowledge to produce some product, or simply exchanging the team of assembled labor with that particular expertise.

The merger process is mediated by search and matching frictions. Firms engage in costly and potentially time consuming search for prospective partners. Upon meeting, the parties bargain over the gains, if any, and choose to consummate some transactions and reject others. After a merger takes place, the acquirer pays a one-time acquisition price to the target, evolves into a new continuing entity with a z determined by the z's of the two pre-merger firms, and continues on in production. Target firms receive the merger payment and exit the market.

To capture key features of the merger market, I extend the classical search paradigm (e.g., Shimer and Smith (2000)) on three dimensions. First, the setup naturally incorporates the large structural changes that occur within firms upon merger - specifically, the firm's z evolves with each match, capturing the dynamic nature of the merger decision not only for current profitability, but also its future merger prospects. Second, I explicitly model repeat acquisition, capturing the prevalence of this phenomenon highlighted in Table 5. Third, because one-time acquirers often themselves become targets, firms search simultaneously on both sides of the market, i.e., as targets and acquirers, although their activities on each side are endogenous and depend on their expected gains from each type of match.[13] Figure 3 illustrates the sequence of steps on the merger market.

Merger technology. Upon entering a meeting, the gains from merger are determined by a "merger technology" to which firms have access. This technology determines the characteristics, i.e., the productivity and profitability, of the post-merger firm, z_m, as a function of the characteristics of the two pre-merger firms, z_a for the acquirer and z_t for the target:

$$z_m = s(z_a, z_t) \tag{2}$$

[13]An alternative setup would be one where firms simply search for partners and their respective roles are determined after observing each other's types. This approach is less consistent with how firms seem to approach the buy/sell decision as detailed, for example, in Boone and Mulherin (2007) and DePamphilis (2009), who describe a process in which firms outline a strategic plan and then search for appropriate partners, so that their roles are well-defined prior to a meeting taking place.

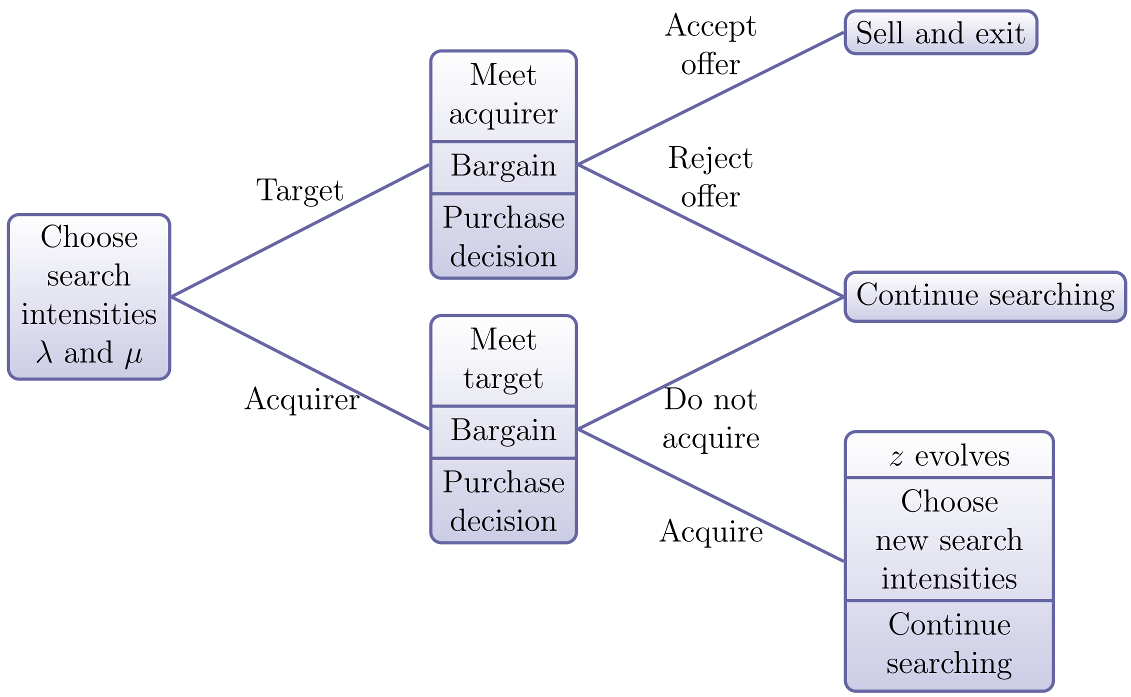


Figure 3: The Timing of Actions on the Merger Market

The shape of the technology plays a pivotal role in determining the gains from merger and empirically disciplining this function is an important piece of my analysis. Section 3.2.1 explores the link with several prevalent theories of merger activity and Section 3.2.2 proposes a parsimonious and flexible aggregator that jointly captures these forces. Section 3.2.3 derives conditions for the technology to lead to sorting, providing strong empirical guidance on its shape and Section 4 builds on those findings for purposes of estimation. For now, I work with the general specification in (2) and defer further discussion until those sections.

Bargaining. The combined gain from merger is the value of the merged entity less the values of the two pre-merger firms as standalone entities:

$$\Sigma\left(z_a, z_t\right) = V\left(z_m\right) - V\left(z_a\right) - V\left(z_t\right) \tag{3}$$

The firms bargain over the gains according to the generalized Nash bargaining protocol, resulting in a commonly agreed upon purchase price $P(z_a, z_t)$. Denoting with β the bargaining power of the acquirer and $1-\beta$ that of the target, the purchase price satisfies

$$P\left(z_a, z_t\right) = V\left(z_t\right) + \left(1-\beta\right)\Sigma\left(z_a, z_t\right)$$

The price depends on both z_a and z_t since the gains depend on the characteristics of both parties and reflects both the outside option of the target, which is to continue as a standalone entity with value $V(z_t)$, as well as the target's share $1-\beta$ of the combined gains. It is straightforward to derive the following expression for the merger premium:

$$\frac{P(z_a, z_t) - V(z_t)}{V(z_t)} = \frac{(1-\beta)\Sigma(z_a, z_t)}{V(z_t)} \tag{4}$$

The premium is equal to the target's share of the combined gains divided by its pre-merger value.

To ease notation going forward, define the individual gains from merger for an acquirer and target of a particular type z partnering with a target of type z_t or acquirer of type z_a as

$$\Sigma_a(z, z_t) = \beta\Sigma(z, z_t), \quad \Sigma_t(z_a, z) = (1-\beta)\Sigma(z_a, z) \tag{5}$$

The literature has often analyzed the gains from merger using percentage gains, specifically, percentage cumulative abnormal stock returns. We can connect the dollar gains in expression (5) to cumulative abnormal returns as

$$CAR_t = \frac{(1-\beta)\Sigma(z_a, z_t)}{V(z_t)}, \quad CAR_a = \frac{\beta\Sigma(z_a, z_t)}{V(z_a)} \quad \Rightarrow \quad \frac{CAR_t}{CAR_a} = \frac{(1-\beta)}{\beta}\frac{V(z_a)}{V(z_t)} \tag{6}$$

The ratio of dollar gains depends only on the ratio of bargaining shares $\frac{1-\beta}{\beta}$, whereas the ratio of abnormal returns depends additionally on the inverse ratio of the pre-merger values. As pointed out by Ahern (2012), expression (6) shows that with significant size differences, percentage abnormal returns may not be indicative of the true split of the gains - a small percentage gain to a large firm may represent a larger dollar gain than does a large percentage gain to a small firm. The bargaining framework naturally gives rise to large disparities in abnormal returns (in favor of targets), even if the true split is more equitable.

Search technology. Firms choose search intensities $\lambda(z)$ of meeting a potential target and $\mu(z)$ of meeting a potential acquirer. To obtain these intensities, the firm incurs costs $C(x)$ for $x = \lambda, \mu$, where $C(\cdot)$ is homogeneous, convex and satisfies $C(0) = 0, C'(\cdot) > 0, C''(\cdot) > 0, \lim_{x\to\infty} C(x) = \infty$. Denote by $dG(z)$ the stationary distribution of firm types in the market, which is kept exogenous at present, but will be made endogenous and in part determined by merger activity in Section 3.3. The aggregate meeting rate, which is the matching function divided by the mass of searching firms, depends on the aggregate search intensities on the two sides of the market and takes the form

$$\min\left\{\int \lambda\left(z\right)dG\left(z\right),\int \mu\left(z\right)dG\left(z\right)\right\}$$

If the number of firms is not equal on the two sides of the market, the short side of the market determines the number of matches and firms on the long side are rationed, i.e., some are left unmatched. The probability of match on the short side of the market is equal to one, and on the long side to the ratio of firms on the short side to the long.

To obtain the matching rates on each side of the market, search intensities must be scaled by these probabilities, which represent the effective meeting rates per unit of search and correspond to the standard notion of market tightness. Define market tightness on the acquirer and target sides of the market as

$$\theta_a=\min\left\{\frac{\int \mu\left(z\right)dG\left(z\right)}{\int \lambda\left(z\right)dG\left(z\right)},1\right\},\quad \theta_t=\min\left\{\frac{\int \lambda\left(z\right)dG\left(z\right)}{\int \mu\left(z\right)dG\left(z\right)},1\right\}$$

The rates at which a type z_a acquirer meets a type z_t target, and in turn, a type z_t target meets a type z_a acquirer are given by

$$\lambda\left(z_a\right)\theta_a\underbrace{\frac{\mu\left(z_t\right)dG\left(z_t\right)}{\int \mu\left(z\right)dG\left(z\right)}}_{\Gamma(z_t)},\quad \mu\left(z_t\right)\theta_t\underbrace{\frac{\lambda\left(z_a\right)dG\left(z_a\right)}{\int \lambda\left(z\right)dG\left(z\right)}}_{\Lambda(z_a)} \tag{7}$$

Examining the first expression, the product of its search intensity $\lambda\left(z_a\right)$ and market tightness θ_a represents the matching rate for this particular acquirer, i.e., the rate at which it meets some candidate target. Next, the presence of a type z_t target in the search market is equal to the product of its search intensity $\mu\left(z_t\right)$ with its density in the firm type distribution $dG\left(z_t\right)$. Conditional on entering a meeting, the probability of an acquirer meeting this particular target is given by the ratio of the target's presence in the search market to the aggregate search intensity on the target side of the market, defined by $\Gamma\left(z_t\right)$. A similar interpretation holds for the rate at which a target of type z_t meets an acquirer of type z_a.

Value functions and decision rules. We can write the value of the firm in a stationary environment as[14]

$$rV\left(z\right)=\max_{\lambda(z),\mu(z)}\quad \pi\left(z\right)-C\left(\lambda\left(z\right)\right)-C\left(\mu\left(z\right)\right)+\lambda\left(z\right)\theta_a\mathbb{E}\left[\Sigma_a\left(z,z_t\right)\right]+\mu\left(z\right)\theta_t\mathbb{E}\left[\Sigma_t\left(z_a,z\right)\right] \tag{8}$$

[14]Conditions to maintain stationarity are detailed in Section 3.3.

where r is the firm's discount rate. The expectations are with respect to the candidate partners the firm may meet.[15] The value of a firm stems from its current flow of profits $\pi(z)$, as well as its prospects on the merger market - expected capital gains less expenditures on search, both as a potential acquirer and target.

The firm makes two types of decisions on the merger market. First, with what intensity to search for merger partners. Optimal search is governed by a pair of first order conditions:

$$C'(\lambda(z)) = \theta_a \mathbb{E}\left[\Sigma_a(z, z_t)\right], \quad C'(\mu(z)) = \theta_t \mathbb{E}\left[\Sigma_t(z_a, z)\right] \tag{9}$$

Firms equate the marginal costs of search to the marginal benefit, where the latter is composed of the incremental probability of a meeting multiplied by the expected gain.

Second, once the firm has met a candidate partner, it must choose whether to consummate the merger or proceed as a standalone entity. This decision is characterized by a pair of acceptance regions representing the set of partners with which the firm is willing to merge. For acquirers, this is the set of targets z_t with which a merger would create positive gains and similarly for targets over the set of acquirers z_a. Formally, define the acceptance regions of acquirers and targets by

$$\Upsilon_a(z) = \{z_t : \Sigma(z, z_t) \geq 0\}, \quad \Upsilon_t(z) = \{z_a : \Sigma(z_a, z) \geq 0\}$$

There is a common acceptance set for acquirers and targets. Any meeting where a merger generates positive combined gains, i.e., where $\Sigma(z_a, z_t) \geq 0$, results in a consummated transaction. In this setup, the sorting of merging firms comes from variation in search intensities and acceptance regions, i.e., matching sets, across firms. These elements provide a sharp link between the nature of merger gains and the strong, though imperfect, sorting patterns observed in the data. The next section conducts a careful investigation of this connection.

3.2 Merger Gains and Merger Patterns

This section uses the model to investigate the link between the merger technology, i.e., the nature of merger gains, and the sorting patterns established in Section 2. First, I analyze the implications of some of the most relevant theories of merger gains. I find that each is consistent with some aspects of observed matching behavior, but none on its own can explain the full set of empirical facts. Next, I propose a flexible Cobb-Douglas technology that can jointly capture these forces and is able to fit the empirical patterns. I use a two-period version of the model to

[15] The firm's expected gains as an acquirer are $\mathbb{E}\left[\Sigma_a(z, z_t)\right] = \int \max\{\Sigma_a(z, z_t), 0\}\, \Gamma(z_t)$, where $\Sigma_a(z, z_t)$ is as defined in (5) and $\Gamma(z_t)$ as in (7), and its expected gains as a target are $\mathbb{E}\left[\Sigma_t(z_a, z)\right] = \int \max\{\Sigma_t(z_a, z), 0\}\, \Lambda(z_a)$.

derive a sharp set of conditions under which this specification can lead to the sorting patterns in the data. These results guide my estimation strategy in Section 4.

3.2.1 Theories of Merger Gains

First, I consider a simple theory of scale efficiencies, where firms merge only for fixed cost savings. Here, the merged firm is simply the sum of the two pre-merger firms, i.e., the merger technology is of the form $z_m = z_a + z_t$, so that there are no particular gains from combining any particular sets of firms. Appendix B.1 proves the following proposition:

Proposition 1. *If the merger technology exhibits only scale efficiencies, i.e.,* $z_m = z_a + z_t$ *and* $c_f > 0$*, then (i) the mean and median differences between acquirers and targets will be zero, (ii) the correlations between acquirers and targets will be zero, and (iii) the median acquirer and median target will be the same as the median firm.*

Merger gains are simply a function of the discounted value of the fixed operating cost, so are constant across all firms and meetings, which leads to random matching. The proposition is then immediate. These predictions stand in contrast to the empirical patterns documented above, which reveal large and systematic differences between acquirers and targets, a high correlation between acquirers and targets and a median acquirer that is considerably larger and more profitable than the median firm.

Next, I consider a theory of purely synergistic mergers as described, for example, in Rhodes-Kropf and Robinson (2008). Here, merger gains are generated from bundling complementarity assets and so by combining firms with similar characteristics, the post-merger firm will be more profitable than the sum of its parts. This notion can be captured by any symmetric and supermodular specification of the merger technology. A natural example is $z_m = A\left(z_a z_t\right)^{\varphi}$, $\varphi > 0$. Under this theory, the following proposition emerges:

Proposition 2. *If the merger technology exhibits pure synergies, i.e., is symmetric and supermodular, then the mean and median differences between acquirers and targets will be zero.*

Key to the result is symmetry - firms equate their effective search intensities on the two sides of the market, implying that there will be no systematic differences between acquirers and targets.[16] Thus, a purely synergistic theory does not explain the substantial and pervasive size

[16]Due to potential asymmetry in bargaining weights, this point is not obvious. All firms actually choose to search more intensively on the side of the market with greater bargaining power. In equilibrium, however, market tightness adjusts to offset the larger degree of aggregate search and equates the effective meeting rates on the two sides of the market.

and profitability differences between acquirers and targets. Section 3.2.3 shows that whether or not such a technology leads to sorting depends on the value of φ.[17]

Third, I consider a version of the q-theory of merger as outlined, for example, in Jovanovic and Rousseau (2002), in which mergers serve as a vehicle for reallocating resources from less productive to more productive firms. Here, merger gains are increasing in the difference between the productivities of the transacting firms, i.e., $\frac{\partial(\Sigma(z_a,z_t))}{\partial(z_a-z_t)} > 0$, since where firms are most different is precisely where there is most room for productivity improvements for the transferred resource, and so gains to be generated. One example is when $\Sigma(z_a, z_t) = V(z_a) - V(z_t)$.[18] Under this theory, we can derive the following proposition:

Proposition 3. *If the merger technology exhibits the q-theory, i.e., gains are increasing in $z_a - z_t$, then (i) low z firms will be overrepresented in the set of targets and high z firms in the set of acquirers, (ii) the median target will be below the median firm and the median acquirer above, and (iii) the highest rate of transaction will occur between low z targets and high z acquirers.*

Merger gains for targets are decreasing in z_t and for acquirers increasing in z_a, so that the lowest z_t's search most intensively and are acquired most rapidly and analogously, higher type z_a's search most intensively and make purchases most rapidly. Then low z_t firms compose the majority of targets and the median z_t will be below the median z and similarly, high z_a firms compose the majority of acquirers. Finally, because the highest and lowest z firms search most intensively and form an acceptable match - indeed, the match that generates the greatest combined gains - they will transact with one another at the highest rate. Thus, q-theory does not explain the facts that target firms are not predominantly low z and that the highest z acquirer and lowest z target almost never transact.[19] Appendix B.1 explores the results of Propositions 1-3 in more detail and, in particular, provides figures comparing each of the implied matching sets to the empirical one.

The q-theory model is one of imitation - by merging, the acquirer allows the target to copy its technology and reaps some share of the gains (which are largest for the highest z_a acquirer and lowest z_t target). This connects to a recent literature on long-run growth through search and imitation, e.g., Perla and Tonetti (2014) and Lucas and Moll (2014). Indeed, Online

[17] E.g., I find that in the presence of repeat acquisition, the example case of $\varphi = 1$ analyzed in Rhodes-Kropf and Robinson (2008) does not generate sorting.

[18] Proposition 3 puts conditions directly on the surplus function, $\Sigma(z_a, z_t)$, and holds whenever the underlying technology satisfies the stated condition. Online Appendix B gives a detailed example of such a technology. In that example, firms have a fixed level of productivity and grow through merger by acquiring additional "segments" (e.g,. plants or products) from the target. I prove that the surplus function from the acquisition of a segment takes exactly this form.

[19] These results are reminiscent of the analysis in Lentz (2010), who proposes a notion of sorting due to differing search intensities across agents, rather than differing matching sets.

Appendix C derives an exact mapping between one version of the q-theory model here and the setup in Lucas and Moll (2014), with the addition of a market for "ideas" - whereas imitation is free in their model and the imitator captures the entirety of the gains, here the firms must both choose to search and transact, and bargain over the surplus. This link suggests that the approach here may be useful for that literature. Specifically, the fact that particular theories of how agents combine knowledge have sharp predictions for how those agents match and sort may help guide future attempts at quantifying those mechanisms using micro-data. For example, given the mapping I derive, the predictions of Proposition 3 are all implications of the growth through imitation framework. Using a similar strategy, it may be possible to use observed sorting patterns (and in the presence of a market, transaction prices) to put more empirical discipline on the process through which the combining and diffusion of knowledge occurs.

3.2.2 A Cobb-Douglas Technology

From here on, I work with a simple yet flexible functional form for the merger technology, namely, an asymmetric Cobb Douglas aggregator of the pre-merger types:

$$z_m = s\left(z_a, z_t\right) = A z_a^{\gamma} z_t^{\nu} \tag{10}$$

It turns out that this specification will be able to fit the data quite well and its simplicity leads to parameters that are easily interpreted and clear intuition for how we can use the data to quantitatively discipline them. The level parameter or "merger productivity," A, captures any autonomous growth from merger independent of the particular characteristics of the transacting firms. The share parameters γ and ν represent the relative weights of the acquirer and target in determining the performance of the post-merger firm. To the extent that these are different, and in particular, that γ is greater than ν, this captures some degree of q-theory, i.e., the acquirer may have more weight than the target in determining the performance of the post-merger entity. Next, the sum of γ and ν plays an important role, as it captures the strength of complementarities, or merger synergies.

An important implication of this specification is that firm productivity is not constant, but evolves with each transaction. With each purchase, the acquiring firm takes on, to some extent, the characteristics of the target. As an example, consider a firm z_1 that purchases, first, firm z_2 to become firm z_3, and then firm z_4 to become z_5. The final entity has productivity

$$z_5 = A z_3^{\gamma} z_4^{\nu} = A\left(A z_1^{\gamma} z_2^{\nu}\right)^{\gamma} z_4^{\nu} = A^{1+\gamma} z_1^{\gamma^2} z_2^{\gamma\nu} z_4^{\nu}$$

The firm evolves into a Cobb-Douglas aggregate of its various components, with the shares determined by the technological parameters and the number of acquisitions (each component depreciates at rate γ upon each new transaction). This formulation has an interesting implication for the acquisition activity of a firm through time. With positive sorting, as a firm makes acquisitions and grows, the types of its partners should also tend to be growing. There is strong support for this prediction in the data - regressions of target profitability or size on the number of prior acquisitions made by the buyer yield positive coefficients that are highly statistically significant (the values are 0.18 and 0.2). This is not necessarily the case under other theories of merger - propositions 1 and 3 feature firm growth through merger, yet matching sets that may remain constant.

3.2.3 Sorting

Propositions 1-3 considered a number of theories of M&A and illustrated that while each is consistent with some aspects of the data, none on its own replicates the full set of empirical patterns. The flexible Cobb-Douglas specification introduced in (10) gives a role for both q-theory and synergistic forces jointly. Asymmetries in the technology will naturally lead to asymmetric matching sets (see Figure 4 for an example) and so generate size/profitability differences between acquirers and targets. In this section, I explore the determinants of sorting. Although Shimer and Smith (2000) may be the most related setup, there are a number of important differences. There, sorting results from complementarities in the production function coupled with option value considerations - once matched, agents cannot continue to match. Here, this second force is not active - because of repeat acquisition, matched firms are able to continue matching. Indeed, a numerical simulation confirms that this difference changes the implications for sorting - under the parameterization of the technology in Figure 1 in Shimer and Smith (2000), sorting occurs in that environment without repeat matching, but the addition of repeat matching results in all agents matching with all others, i.e., matching is random.

Let $l(z) = \min\{z_a | z_a \in \Upsilon_t(z)\}$ and $u(z) = \max\{z_a | z_a \in \Upsilon_t(z)\}$ denote the lowest and highest type acquirer acceptable to a target of type z. There is positive sorting if (i) $l(z)$ and $u(z)$ are nondecreasing in z, (ii) either $l(z)$ or $u(z)$ is strictly increasing at some z and (iii) $\Upsilon_t(z)$ is nonempty and convex for all $z \geq z_{min}$ where z_{min} is the smallest z such that $z_{min} \in \Upsilon_t(z_{min})$. Under positive sorting, (i) matching sets are characterized by weakly increasing upper and lower bound functions, (ii) there is at least least one pair $z_1 < z_2$ such that one of the bounds is higher for z_2 than z_1, which implies that $z_{max} \notin \Upsilon_t(z_{min})$ and so a matching set consisting of the entire space of z does not constitute positive sorting, and (iii) these properties hold only for the set of z's that match with their own type.[20]

[20] This definition is an adaptation of Atakan (2006). Shimer and Smith (2000) prove it is equivalent to the "lattice" definition they employ if matching sets are closed and nonempty. The presence of sorting is cast in terms

Deriving conditions that guarantee sorting in the infinite-horizon model is challenging. However, it turns out we can gain intuition from a simpler two-period example that yields sharp conditions for sorting. Numerical simulations suggest that these conditions remain necessary in the full infinite-horizon setting. The setup builds closely on that in Eeckhout and Kircher (2011). In period 1, all firms are randomly matched and choose whether to consummate a merger with their partner or not. In period 2, firms are matched in a frictionless matching market. The payment to the target in period 1 is determined by Nash bargaining. For simplicity, I focus on the symmetric case where $\gamma = \nu = \varphi$, $\beta = \frac{1}{2}$ and $A = 1$, so that $z_m = (z_a z_t)^{\varphi}$.[21] Following Eeckhout and Kircher (2011), Appendix B.2 shows that in period 2, matching is perfectly assortative, so that $z_m = z^{2\varphi}$ for type z and each firm receives $\frac{1}{2}z^{2\varphi}$.

Consider two firms z_a and z_t who are matched in period 1. If they reject the match, firm z_a accrues profits z_a in the first period and $\frac{1}{2}z_a^{2\varphi}$ in the second, and similarly for firm z_t. If they merge, the merged firm accrues profits $z_m = (z_a z_t)^{\varphi}$ in the first period and, in a key departure from Eeckhout and Kircher (2011), rematches in the second period, when it receives $\frac{1}{2}z_m^{2\varphi} = \frac{1}{2}(z_a z_t)^{2\varphi^2}$. The total surplus from a match in the first period is

$$\Sigma(z_a, z_t) = (z_a z_t)^{\varphi} - z_a - z_t + \frac{1}{2}(z_a z_t)^{2\varphi^2} - \frac{1}{2}z_a^{2\varphi} - \frac{1}{2}z_t^{2\varphi} \tag{11}$$

and first period matches will be consummated if this is positive. The first three terms correspond to the one-period surplus, i.e., how much higher are current merged profits than those of the two standalone firms and the second three terms take into account the dynamic considerations of (re)matching in the following period.

Appendix B.2 proves sorting obtains if and only if $\varphi \in \left(\frac{1}{2}, 1\right)$. Intuitively, for $\varphi \leq \frac{1}{2}$, merged profits never exceed standalone profits and the matching set is empty. For $\varphi \geq 1$, all mergers lead to static profit gains and given that the firms will rematch, all mergers are consummated and matching is random. This result can be understood as two conditions - (i) complementarities (or "synergies") must be sufficiently strong, such that $\varphi > \frac{1}{2}$, and (ii) the technology must exhibit sufficient curvature with respect to each argument, such that the elasticity satisfies $\frac{\partial \log z_m}{\partial \log z_i} < 1$ for $i =, a, t$.[22] In the more general asymmetric case, the conditions correspond to (i) $\gamma + \nu > 1$ and (ii) $\gamma < 1, \nu < 1$.[23]

---

of matching set variation across firms. Lentz (2010) provides an alternative notion in a related environment featuring on-the-job search and no option value on the part of firms. He derives conditions on the production function such that more intensive search by high skill workers leads to sorting in a stochastic dominance sense, despite all agents having a common matching set. Section 5.2 explores this distinction further by quantifying the contribution of M&A from each of these margins individually.

[21] All results go through for reasonable values of A.

[22] For the Cobb-Douglas technology, this implies negative second derivatives and so a standard notion of concavity in each argument of $s(\cdot)$.

[23] Numerical simulations suggest these conditions remain necessary in the asymmetric case.

The lower bound on curvature is a novel result to this setting - even when the usual conditions on complementarities are satisfied, sorting fails when the technology does not exhibit curvature with respect to each argument. This is due to the presence of repeat acquisition. To see this, consider the alternative case where if firms merge in period 1, they cannot do so in period 2. Appendix B.2 proves that while $\varphi > \frac{1}{2}$ remains a necessary condition, the bound on curvature is no longer needed, i.e., there will be sorting even for $\varphi = 1$ (and numerical simulations confirm sorting for values greater than one). With repeat acquisition and insufficient curvature, surplus is always positive and firms will merge with any potential partner; without repeat acquisition, firms consider the option value of foregone future matches and reject types that are too high or too low, even in the absence of curvature.

The curvature condition can be related to the presence of heterogeneous standalone profits (that can exceed merged profits). For example, the model in Shimer and Smith (2000) assumes standalone profits are zero. We can restate this difference in terms of the technology only - it is straightforward to prove that the Shimer and Smith (2000) model is isomorphic to one in which each agent's standalone profit is equal to its type and the technology is $z_1+z_2+s(z_1, z_2)$. Clearly, in this formulation, static profits are always higher after merging and so with repeat matching, all firms merge, leading to random matching (proof in Appendix B.2). The technology does not exhibit sufficient curvature.

The lower bound on complementarities is also related to the presence of standalone profits. Even when the technology satisfies standard conditions on complementarities, if they are not sufficiently strong, standalone profits always exceed merged profits and no firms will merge. The Shimer and Smith (2000) model corresponds to one in which there is no repeat matching and the technology takes the form $z_1+z_2+s(z_1, z_2)$, so merged profits always exceed standalone profits. Appendix B.2 proves that in this setup, sorting obtains for any value of $\varphi > 0$, as in their paper, and the lower bound on complementarities is no longer necessary.

In sum, the two-period example points to sharp conditions for sorting in the presence of repeat acquisition - both sufficient complementarities and curvature. Simulations suggest these conditions are necessary in the infinite-horizon model as well. Specifically, the simulations show precisely the same results as here - no mergers if $\varphi \leq \frac{1}{2}$, all matches result in merger if $\varphi \geq 1$ and positive sorting for $\varphi \in \left(\frac{1}{2}, 1\right)$. In addition to revealing the drivers of sorting, these results are useful because they put further discipline on the admissible shape of the technology, e.g., the parameters must satisfy the stricter conditions for sorting in this environment.

3.3 General Equilibrium

Before turning to the estimation, I close the model in general equilibrium.

Preferences and production. A measure L of identical households inelastically supply labor to firms and value consumption C of a single homogeneous good. Denote by ρ the household's rate of time discount, which is equal to the real interest rate.

Firms are competitive and produce output using labor according to a decreasing returns to scale production function $q(z) = z^{1-\alpha} l(z)^{\alpha}$ as in the Lucas (1978) span-of-control model. Profit maximization gives revenue, employment, and variable profits as

$$r(z) = \frac{1}{1-\alpha}\Pi z, \quad l(z) = \frac{1}{1-\alpha}\frac{\alpha}{w}\Pi z, \quad \pi(z) = \Pi z \tag{12}$$

where $\Pi = (1-\alpha)\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}$. Firm product market outcomes - most importantly, variable profits - are proportional to z as assumed thus far and depend on parameters and aggregate variables that are common across all firms, here the span of control parameter α and the wage w.[24]

Entry and exit. Firms enter and exit the economy. Incumbent firms are subject to an exogenous exit shock that arrives at rate δ, common across all firms. Additionally, exit comes when a firm is acquired. The rate of exit for an incumbent firm of type z is equal to

$$\delta + \mu(z)\,\theta_t \int \Phi\left(\Sigma_t(z_a, z)\right)\Lambda(z_a)$$

where the second term is the rate at which this firm type is acquired, which is the product of its meeting rate as a target and the conditional probability that a transaction is consummated. This last is the integral over the set of acquirers with which a merger generates positive gains, weighted by the presence of firm types on the acquiring side of the market as set out in (7). $\Phi(\cdot)$ denotes the indicator function equal to 1 if its argument is greater than or equal to zero, else equal to zero. Despite the common exit shock δ, effective exit rates will vary systematically across incumbent firm types to the extent that the rate of being acquired does.

There is a large pool of ex-ante identical potential entrants. To enter, entrepreneurs must expend c_e units of goods to obtain an initial z draw from an exogenous distribution $F(z), z \in (z_{\min}, \infty)$ with associated density function $dF(z)$. Free entry requires that in an equilibrium with positive entry - the case I focus on - the expected value of entry equal the cost:

$$\int V(z)\,dF(z) = c_e \tag{13}$$

Once an entrant realizes its initial z, it may enter the market and begin operations or exit

[24]The framework also accommodates a setup with differentiated goods, CES demand and monopolistic competition as in Melitz (2003) or Bertrand competition with limit pricing as in Bernard et al. (2003). The key is that firm profitability is linear in some measure of firm type.

immediately, and will only choose to enter if it garners positive value from doing so. The presence of the fixed operating cost c_f implies that some entrepreneurs may draw a low enough z so that the value from entering is negative - in other words, there is selection. The entry decision is determined by a threshold $\hat{z}$ defined implicitly where the value of entry is equal to zero, i.e., $V(\hat{z}) = 0$. Entrepreneurs drawing $z \geq \hat{z}$ will choose to enter while those with $z < \hat{z}$ will exit immediately. Using expression (8), the threshold entrant must satisfy

$$\pi(\hat{z}) = -\left\{\lambda(\hat{z})\,\theta_a \mathbb{E}\left[\Sigma_a(\hat{z}, z_t)\right] + \mu(\hat{z})\,\theta_t \mathbb{E}\left[\Sigma_t(z_a, \hat{z})\right] - C(\lambda(\hat{z})) - C(\mu(\hat{z}))\right\} \tag{14}$$

The threshold is set where the firm's net profit equals the negative of the value it obtains from its M&A prospects, expected gains less costs. This value must be nonnegative, or the firm would choose not to participate in the M&A market at all. Equation (14) shows that there will be a set of entrepreneurs that enter - even while accruing negative profits - only hoping to take advantage of some value-increasing opportunity on the merger market. Without the possibility of merger, the threshold would be set where profits are exactly zero. The expression shows that all else equal, by directly increasing the value from entry, M&A drives down the entry threshold, i.e., induces less productive firms to enter than would without M&A, and reduces the selection of new entrepreneurs.

In general equilibrium, however, it is not necessarily the case that selection worsens due to M&A. The entry decision also depends upon the level of firm profitability, i.e., Π, which is a function of the economic aggregates, specifically, the wage. To the extent that M&A leads to productivity gains, this will increase wages (by increasing labor demand), reduce firm profitability and make successful entry more difficult, leading to increased selection. Thus, the effect of M&A on the entry decisions of new entrepreneurs is ambiguous and depends on the rate at which profitability falls due to gains in aggregate efficiency versus the rate at which the potential gains from merger add to the marginal firm's value, i.e., whether the left hand side of (14) falls faster or slower than the absolute value of the right hand side increases.

Aggregation. The model economy aggregates in a simple fashion. Aggregate output, productivity, and the real price level (the inverse of the wage) satisfy

$$Y = \left(M\bar{Z}\right)^{1-\alpha} L^{\alpha}, \quad TFP = \left(M\bar{Z}\right)^{1-\alpha}, \quad P = \frac{1}{\alpha}\left(M\bar{Z}\right)^{\alpha-1} L^{\alpha} \tag{15}$$

where M denotes the mass of operating firms and $\overline{Z} = \int_{\hat{z}}^{\infty} z dG(z)$ an index of average firm productivity. As is standard in this class of model, once M and $\overline{Z}$ are determined, the economy behaves as one with a representative firm with productivity equal to TFP as defined in (15). M&A in part endogenizes the components of TFP and affects aggregate outcomes through two key statistics - the mass of operating firms, M, and their average productivity, $\overline{Z}$. The latter depends on (1) any direct productivity gains generated through merger and the associated effect on the distribution of firms, $dG(z)$ (i.e., the "intensive margin"), and (2) the effects on the threshold productivity level, $\hat{z}$, i.e., the degree of selection of new entrants. The change in the mass of firms depends on whether any additional entry of new entrepreneurs is sufficient to offset the removal of firms through acquisition.

Equilibrium. In a stationary equilibrium, the aggregate variables remain constant through time, implying that the inflows and outflows of firms in the market must balance for all types. The stationary conditions for each type $z \geq \hat{z}$ take the form

$$\begin{aligned} & M\int \lambda\left(z_a\right)\theta_a\Phi\left(\Sigma\left(z_a, s^{-1}\left(z, z_a\right)\right)\right)\Gamma\left(z_t\right)dG\left(z_a\right) + M_e dF\left(z\right) \\ = \quad & \lambda\left(z\right)\theta_a M dG\left(z\right)\int \Phi\left(\Sigma\left(z, z_t\right)\right)\Gamma\left(z_t\right) \\ + \quad & \mu\left(z\right)\theta_t M dG\left(z\right)\int \Phi\left(\Sigma\left(z_a, z\right)\right)\Lambda\left(z_a\right) + \delta M dG\left(z\right) \end{aligned} \tag{16}$$

where $s^{-1}(z, z_a) = \{z_t : s(z_a, z_t) = z\}$ denotes the inverse of the merger technology defined in (10) and M_e the mass of new entrants. For each type, firms flow in as the continuing entity from merger and through entry by new entrepreneurs. Firms flow out through participation in a merger, either as an acquirer or target, and through exogenous exit. Integrating (16) gives a related condition equating the flows of firms into and out of the economy.

There are two feasibility constraints. First, labor market clearing requires that $L = L_p$, where L_p denotes the total labor demand of firms. Goods market clearing requires

$$Y = C + Y_s + Y_f + Y_e \tag{17}$$

where

$$Y_s = M\left[\int C_\lambda\left(\lambda\left(z\right)\right)dG\left(z\right) + \int C_\mu\left(\mu\left(z\right)\right)dG\left(z\right)\right]$$

denotes the total resources devoted to search activities on the merger market, $Y_f = Mc_f$ to the fixed costs of production, and $Y_e = M_e c_e$ to new firm creation.

A *stationary search equilibrium* consists of (1) aggregate variables $\{Y, P, C, M, M_e, dG(z)\}$, (2) firm profits, entry threshold, and values $\{\pi(z), \hat{z}, V(z)\}$ and (3) firm search intensities and acceptance sets $\{\lambda(z), \mu(z), \Upsilon_a(z), \Upsilon_t(z)\}$, such that (a) firms maximize expected discounted profits, (b) the labor and goods markets clear and (c) the evolution of firm types is consistent with the stationary conditions.

4 Estimation and Model Fit

In this section, I describe my estimation approach and the resulting parameter values. I assess the fit of the estimated model on the empirical patterns in Section 2.

4.1 Empirical Strategy

First, I make several normalizations and assign values to a small number of parameters where identification is not dependent on the model. A time period is assumed to be one year. I normalize the mass of consumers L to be 1 and the sunk cost of entry c_e to the same. The real interest rate ρ is set to 5%. I set the exit shock δ to 0.04 to match the exit rate of publicly listed firms such as those in Compustat.[25] I set the entry distribution $dF(z)$ so that the endogenous distribution $dG(z)$ is Pareto with shape parameter ξ and minimum value $\hat{z}$.[26] A large number of studies find that the firm-size distribution closely approximates a Pareto with a shape parameter close to one.[27] Following Atkeson and Burstein (2010), I set the shape parameter $\xi = 1.2$. I set c_f such that $\hat{z}$ is normalized to one.

Identification of the merger technology builds closely on the link between merger gains and sorting patterns described in Section 3.2. For example, consider a simple one-shot version of the model. Firms are randomly matched and decide whether to consummate the merger or not. If not, each firm operates separately and receives profits Πz. If they do, the firms combine and receive profits Πz_m. Owners of the acquiring firm are entitled to a share β of these profits and of the target to a share $1-\beta$. For the merger to go through, each firm's share of the new profit stream must be greater than its standalone profits. Using the technology in (10) gives closed form expressions for the upper and lower bounds on the matching set:

$$\begin{aligned}
\beta z_a^{\gamma} z_t^{\nu} &\geq z_a \Rightarrow \log z_a \leq \frac{1}{1-\gamma}\log(\beta A) + \frac{\nu}{1-\gamma}\log z_t \\
(1-\beta) z_a^{\gamma} z_t^{\nu} &\geq z_t \Rightarrow \log z_a \geq -\frac{1}{\gamma}\log((1-\beta)A) + \frac{1-\nu}{\gamma}\log z_t
\end{aligned}$$

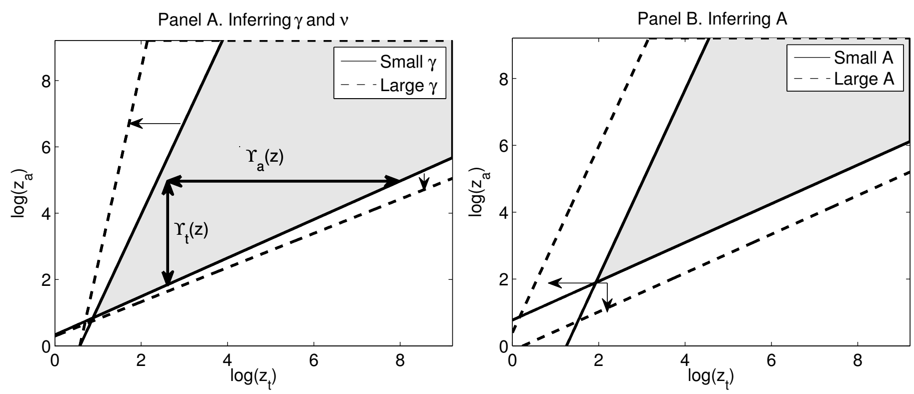
Appendix C.1 proves the conditions for sorting derived in Section 3.2 are necessary and sufficient here as well (with repeat acquisition, the dynamic matching problem resembles the static one). Assuming these conditions hold, the shaded region in Panel A of Figure 4 plots

[25] To calculate this value, I use data from the Center for Research in Security Prices (CRSP) database, which provides information on exit (delisting) of publicly traded firms. I exclude delisting codes 200-299, which are related to M&A and so are not true exit. I use the annual average over the period 1980-2009.

[26] Specifically, the estimation imposes the specified distribution for $dG(z)$, solves for firm search and matching decisions and uses (16) to infer the primitive distribution $dF(z)$. This gives the same result as looping over the primitive distribution at considerably less computational cost. Details are in Online Appendix G.

[27] See, for example, Axtell (2001) and references thereto.

an example matching set, where any meeting of firms between the solid lines results in a consummated merger. The figure shows that an increase in γ rotates the bounds outward to the dashed lines, expanding the matching set to include a greater number of less productive firms. A similar rotation occurs when increasing ν. Following this intuition, I estimate γ and ν to find the bounds that best approximate the matching observed in the data. Specifically, as γ and ν in large part determine the rates of search and matching across the set of firms in the market and so the rate of transaction for each firm type, I estimate their values to target the median deviation of acquirer and target profits from the median in their industries.



Notes: The shaded area in Panel A displays a theoretical matching set and the dashed lines the boundaries of the new matching set for an increase in the weight of the acquirer in the merger technology, γ. Horizontal segments represent the acceptance regions of acquirers, $\Upsilon_a(z)$, and vertical segments the acceptance region of targets, $\Upsilon_t(z)$. Panel B similarly displays the effect of changes in the level parameter A.

Figure 4: Intuitive Identification of Merger Gains

The example in Panel A excludes a number of matches between small firms that fall to the southwest of the shaded region. Panel B shows that since A appears only in the intercepts of the matching bands, an increase in A results in a parallel shift of the bands (from the shaded region to the dashed lines), increasing the number of small transactions. I estimate A to match the number of mergers among small firms, specifically, the percentage of targets that fall in the bottom decile of the firm size distribution. In Figure 9 in Appendix C.1, I provide a numerical analog of these identification arguments using the fully estimated model. Specifically, I vary the three parameters of the technology around their estimated values (holding the other parameters fixed) and plot the resulting matching sets. The intuition from this example goes through almost exactly.

To pin down the Nash bargaining weight β, equation (4) directly relates the merger premium to the bargaining shares, and in particular, shows that the premium reflects both the combined gains from merger, and the target's share. Once we know the gains from merger, we can use (4) in conjunction with data on the premium to infer the target's bargaining share $1 - \beta$. I estimate the bargaining parameter β to target the mean merger premium.

The search cost function takes the form[28]

$$C(x) = \frac{B}{\eta}x^{\eta}, \quad \eta > 1, \quad \text{for } x = \lambda, \mu$$

I estimate the scale parameter B so that the aggregate merger rate in the model matches that in the data. From the combined SDC and Compustat data, I find that about 3.7% percent of Compustat firms are acquired annually over the sample period. This value is in line with evidence from Maksimovic and Phillips (2001) who report that an annual average of 3.9% of large manufacturing plants in the US changed ownership in the Longitudinal Research Database over the period 1974-1992. Intuitively, having used the observed matching patterns to infer the gains from merger, I am essentially finding the costs that reconcile these gains with the empirical merger rate. The curvature parameter, η, influences the distribution of search intensities across firms. A high value of η implies a fast-increasing cost of search and will tend to spread out search across the range of firms. A low value of η implies the opposite, causing search intensities to be more concentrated within those firms with the most to gain from merger. I estimate η to match the dispersion in target types, measured by the coefficient of variation in target profits $\frac{std(\pi(z_t))}{mean(\pi(z_t))}$.

I estimate the model using a simulated method of moments approach (details are in Appendix G). Formally, I search over the parameter vector $(\gamma, \nu, A, \beta, B, \eta)$ to find the combination that minimizes a distance criterion between the simulated and empirical moments. The parameters are estimated jointly and there are not analytic expressions for the moments of interest. However, in a structural model such as the one here, it is useful to have a sense of the relationship between moments and parameters. Towards this end, Appendix C.1 numerically demonstrates the relationship between each moment and parameter and confirms the intuitions laid out above, specifically, that each moment is sensitive to the corresponding parameter.

Table 6 reports the estimated parameter values, as well as the empirical and simulated moments. The model matches the six target moments closely. Standard errors are bootstrapped and are generally small, a not surprising result given the number of observations in the sample.

[28] I assume an equal level of search costs on the two sides of the market. Appendix C.2 proves that this is without loss of generality for any disparity in search costs so long as $\theta_t < 1$, which, at the estimated parameters, is the case when the level of target search costs is not more than eight times higher than that of acquirers. For any level of target costs in this range, the proof shows that the parameter estimates and quantitative exercises are invariant to the level of those costs.

Table 6: Merger Market Parameter Estimates

Parameter	Estimate	Standard Error	Target Moment	Model	Data
γ	0.91	0.0048	Median acquirer profitability	0.57	0.57
ν	0.53	0.0088	Median target profitability	0.00	-0.01
A	1.05	0.0115	Share of targets in lowest decile	0.08	0.10
β	0.49	0.0065	Mean merger premium	0.47	0.47
η	14.25	0.1274	Dispersion of target profitability	3.92	3.92
B	3.44	0.0463	Acquisition rate (x 10^{12})	0.04	0.04

Notes: Table reports estimation results of merger market parameters: point estimates and standard errors, target moments and their values in the data and estimated model. Median acquirer and target profitability are the median deviation of log profits from the median in their industry and the measure of dispersion is the coefficient of variation of log profitability, as described in the text. Standard errors are bootstrapped.

4.2 The Gains and Split from Merger

The share parameters of acquirers and targets in the merger technology, γ and ν, are estimated to be 0.91 and 0.53, respectively. There are three key takeaways: first, this represents a significant departure from symmetry - the acquirer holds substantially more weight than the target in determining the performance of the post-merger entity. The technology exhibits some degree of q-theory, i.e., some capability of acquiring firms to improve on the performance of their targets. Second, the sum of γ and ν is about 1.44, significantly greater than one - sizable synergies can be realized from combining compatible sets of firms. Thus, the data point to important roles for both forces in leading to merger gains.[29] Third, each share parameter is less than one - if the potential match is poor, merging would lead to losses in productivity and profitability compared to remaining as standalone entities (put another way, if the pre-merger target is too unproductive relative to the acquirer, the prospect of productivity gains on the transferred asset is trumped by losses suffered on existing assets). These values satisfy the conditions derived in Section 3.2.3, i.e., (i) $\gamma + \nu > 1$ and (ii) $\gamma < 1,\ \nu < 1$.

The bargaining parameter β is estimated to be 0.49 - the gains from merger are split almost equally. Recall that the merger premium reflects both target bargaining power and the size of the combined gains. Given the estimates of merger gains, the results suggest that bargaining shares must be relatively balanced to imply premia on the order of magnitude observed in the data. Section 4.3 revisits the model's quantitative implications for the gains and split from

[29] Related findings are in Maksimovic and Phillips (2001), who find that the productivity of transferred assets generally improves following an ownership change. Schoar (2002) finds similarly, but additionally that the productivity of acquirer assets generally falls. Bhagat et al. (2005) find that takeover gains tend to be higher when acquirer and target are closer in size, but conditional on the size of the target, are increasing in the size of the acquirer. Their results are exactly consistent with the estimated merger technology here, and they conclude with a similar interpretation that "these findings are consistent with the importance of both synergies and target-specific improvements such as removal of bad management."

merger in light of some of the salient findings of the empirical corporate finance literature.

4.3 Non-Targeted Moments

The method of moments approach infers the parameters of the model using an informative set of moments that capture key aspects of the data. This section demonstrates that the parameter estimates are consistent with the broader set of empirical patterns documented in Section 2 - specifically, the joint and marginal distributions of transacting firms - as well as with existing empirical findings on the gains and split from merger.

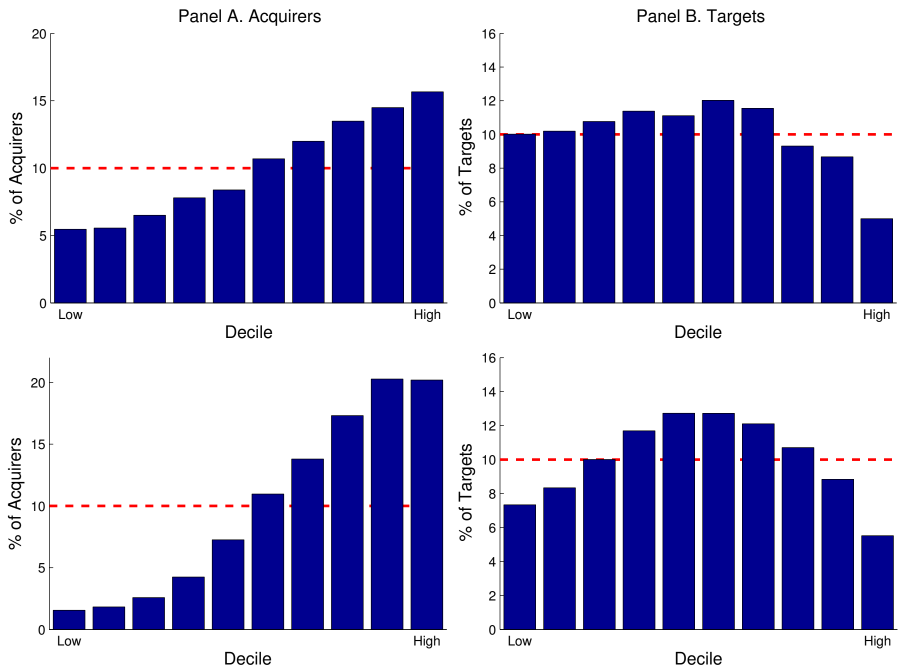
Joint and marginal distributions. Table 7 shows that the model closely replicates the two main features of the empirical joint distribution - strong positive sorting alongside prevalent profitability and size differences. The first four rows compare the correlations of acquirer and target profits, sales, employment, and market values from the estimated model and the data. The last row reports the fraction of transactions in which the size and profitability of the acquirer exceeds that of the target (the data value is an average over the metrics in Table 2). On both dimensions, the model predicts a degree of sorting very close to that in the data.

Table 7: Sorting in M&A: Model and Data

Moment	Model	Data
Correlations:		
Profits	0.58	0.63
Sales	0.58	0.62
Employment	0.58	0.58
Market value	0.71	0.64
Share of transactions with $z_a > z_t$	0.86	0.90

Notes: Table reports log correlations of acquirer and target characteristics in the data and from the estimated model, and the proportion of transactions in which the acquirer exceeds the target.

Figure 5 shows that the model matches the marginal distributions of both acquirers and targets. The top row replicates Figure 2 in showing the empirical distributions and the bottom row the corresponding distributions generated from the model. Turning to acquirers in Panel A, the model generates the monotonic pattern observed in the data, with acquirers overrepresented in the top 5 deciles and underrepresented in the bottom 5. The model somewhat underpredicts the share of acquirers at the bottom and overpredicts the share at the top, but overall fits the empirical distribution closely and captures its key features. Similarly, Panel B shows that the model generates the overrepresentation of targets in the middle deciles, slightly underpredicts their share in the lowest deciles and fits closely the sharp drop-off in the top decile.



Notes: Figure displays the proportion of transacting firms that fall into each decile of the firm size distribution (measured by profitability) in the data and from the estimated model.

Figure 5: Marginal Distributions of Transacting Firms: Data (top) vs Model (bottom)

Implications for asset prices. To measure the gains from merger, much of the empirical literature relies on an event study approach using cumulative abnormal stock market returns (CAR) at the time the merger is announced. The model equivalent to the CAR is given by $CAR = \frac{\Sigma(z_a,z_t)}{V(z_a)+V(z_t)}$ where the numerator represents the combined dollar gains and the denominator the combined pre-merger market values. Using the following representation

$$CAR = \frac{prem}{1-\beta}\frac{V\left(z_t\right)}{V\left(z_a\right)+V\left(z_t\right)}$$

along with a mean premium of 47%, β of 0.49, and a mean ratio of $V\left(z_a\right)$ to $V\left(z_t\right)$ of 10 (e.g., Table 2) gives a model-implied mean CAR of about 8% (averaging over the model-implied distribution of matches gives a similar value, 7.2%).

Typical estimates of the mean combined CAR are between 2% and 3%.[30] However, the literature has pointed out several factors that may confound identification of merger gains using the CAR, specifically, that put a downward bias in the measured gains to acquirers. Important examples include (1) price changes due to contemporaneous financial transactions, (2) truncation of the event window so that the estimates reflect uncertainty regarding the final

[30]See, for example, Andrade et al. (2001), Betton et al. (2008), and Masulis et al. (2008).

outcome of the bid and (3) negative price reactions to the revelation of private information regarding the acquirer's standalone prospects.[31] After controlling for these factors, estimates of the average combined gains range anywhere from 4% to over 17%.[32] The model's prediction of about 8% is slightly below the midpoint of this range.

The literature overwhelmingly finds that targets capture the lion's share of the gains, an inference based on the fact that the abnormal returns to the target are generally substantial, while those to the acquirer are generally much smaller and actually negative in a significant number of transactions, with an overall mean of about 1%.[33] In contrast, I find the split to be almost equal. Equation (6) shows why these results are consistent - the model CAR ratio is $\frac{CAR_t}{CAR_a} = \frac{(1-\beta)}{\beta}\frac{V(z_a)}{V(z_t)}$. In the presence of size differences, disparities in percentage abnormal returns are not indicative of disparities in the true split of dollar gains. Quantifying this relationship using the estimated value of β, the results imply CAR_t should exceed CAR_a by a factor of 10, despite the fact that the true split is almost equal.

Alternative approaches. Guided by the link between merger gains and matching patterns laid out in Section 3.2, my estimation strategy relies primarily on matching key moments capturing the ex-ante characteristics of merging firms. Given that both pre- and post-merger performance measures are observable, one alternative would be to estimate the parameters of the technology by investigating the relationship between the post-merger firm, z_m, and the pre-merger firms, z_a and z_t. There are a number of factors that make this approach challenging. For example, Li and Prabhala (2007) point to the decision to merge as a leading example of sample selection concerns in corporate finance. In Appendix C.3, I add a match-specific component to the technology, potentially observed by the firm, but not the econometrician. This leads to a sample selection model and a downward bias on the coefficients obtained from pre/post-merger regressions. In contrast, I show that my strategy exhibits a robustness to this issue. I simulate data from an environment featuring match-specific shocks and endogenous selection and estimate the parameters of the (now misspecified) technology using both approaches. The results show a significant downward bias in the estimates from the regression-based approach,

[31]Braguinsky and Jovanovic (2004) develop a theoretical environment in which mergers always improve firm performance, yet the revelation of new information to shareholders upon merger announcement causes the acquirer's value to fall, as can the combined values of the acquirer and target.

[32]For example, Andrade et al. (2001) find that after removing stock-financed transactions to account for new equity issues, the mean CAR is between 4% and 5%. Bhagat et al. (2005) find combined gains of about 7% after adjusting for truncation bias. Recent work measuring the "revelation bias" finds a mean CAR ranging from 9% in Wang (2012) to 13% in Bhagat et al. (2005) and as high as over 17% in Masulis et al. (2008).

[33]See, for example, Andrade et al. (2001), Betton et al. (2008), and Moeller et al. (2004). Addressing the significant fraction of acquiring firms that realize negative returns is largely the point of the literature assessing the role of the revelation effect in biasing these estimates. Moreover, the adjustment suggested by Andrade et al. (2001) to account for associated new equity issues is likely to play a role here as well.

but much smaller effects when using the strategy I employ.[34]

Although relating the coefficients from the pre/post-merger regressions to the parameters of the technology is difficult, it can still be useful to get a sense of the relationships that they uncover. A number of papers have estimated related regressions and typically find positive and significant effects of M&A activity on subsequent firm performance, recent examples of which include Maksimovic and Phillips (2001), Maksimovic et al. (2011), Li (2013) and Braguinsky et al. (2015). In Appendix C.3, I confirm that these findings hold in my data as well. Specifically, I follow a similar approach to these papers and estimate regressions of firm performance on an acquisition indicator variable, controlling for prior performance and industry-year fixed-effects. In line with that literature, the results point to statistically and economically significant post-merger performance gains.

Hagedorn et al. (2017) propose a related approach to non-parametrically identify the technology in a labor market context using pre-match values and wages. Online Appendix D shows that applying their methodology in an environment with repeat matching and endogenous search intensities is challenging. In the presence of these additional elements, we cannot derive an expression for the output of a match in terms of pre-match values and the transaction price alone, a key step in their method.

Random search. The assumption of random search plays an important role in my analysis. One implication of the random search environment is as follows: consider two intervals of acquiring firms $[z_{a,1}, z_{a,2}]$ and $[z_{a,3}, z_{a,4}]$, where $z_{a,3} > z_{a,2}$. For an appropriate choice of these intervals, there will be a common set of acceptable targets, $[z_{t,1}, z_{t,2}]$. Random search implies that the distribution of matches over the interval $[z_{t,1}, z_{t,2}]$ should be the same for acquirers in sets $[z_{a,1}, z_{a,2}]$ and $[z_{a,3}, z_{a,4}]$.

In Appendix C.4, I investigate this prediction in detail. I describe one way to systematically form these intervals by allocating firms into two-way sorted deciles (i.e., for each acquirer decile group, sort targets into their deciles). I conduct a Kolmogorov-Smirnov (KS) test of the equality of distributions to ascertain whether the distribution of targets within the intervals defined by those deciles is the same for acquirers in adjacent deciles. For example, consider group (1,6), the sixth decile of targets purchased by the lowest decile of acquirers. The lower and upper bounds of this group are targets with \$1.8 million and \$2.3 million in profits, respectively. This defines one target interval. I test whether for targets in this range, the distribution is the same for acquirers in decile one and decile two. I perform analogous tests for all intervals defined by target deciles 6-10. This gives a total of 45 tests. Of the 45, the KS-test fails to reject the null

[34]Hagedorn et al. (2017) demonstrate a similar robustness to unobserved match-specific shocks using an approach relying on ex-ante sorting and prices (wages).

of equal distributions in 43 of them.

5 The Aggregate Implications of M&A

In this section, I use the estimated model to evaluate the implications of M&A for aggregate outcomes. Expression (15) provides a simple link between M&A and industrial outcomes, e.g., aggregate productivity/output. Using that equation, write the percentage change in output when moving to a counterfactual economy as

$$\Delta \log Y = (1-\alpha)\,\Delta \log M + (1-\alpha)\,\Delta \log \overline{Z} \tag{18}$$

where $\Delta \log X = \log X_{cf} - \log X$ denotes the percentage (log) change in a variable when moving from the benchmark economy (no subscript) to the counterfactual (subscript cf). Equation (18) decomposes the change in output into changes in the mass of firms ("entry") and average firm productivity, each with an elasticity of $1-\alpha$. If $\hat{z}_{cf} < \hat{z}$, we can decompose changes in $\overline{Z}$ as

$$\Delta \log \overline{Z} = \underbrace{\left[\log \int_{\hat{z}}^{\infty} z d\tilde{G}_{cf}(z) - \log \int_{\hat{z}}^{\infty} z dG(z)\right]}_{\text{intensive margin}} + \underbrace{\left[\log \int_{\hat{z}_{cf}}^{\infty} z dG_{cf}(z) - \log \int_{\hat{z}}^{\infty} z d\tilde{G}_{cf}(z)\right]}_{\text{selection}} \tag{19}$$

where $d\tilde{G}_{cf}(z) = \frac{dG_{cf}(z)}{1-G_{cf}(\hat{z})}$.[35] The first term captures the change on the intensive margin - fixing the threshold productivity level, what is the effect of M&A through the changing distribution of firm types. The second term captures selection - fixing the distribution, what is the effect of M&A due to changes in the threshold productivity level. Scaling these terms by $1-\alpha$ gives their contributions to changes in output.

Consumer outcomes depend additionally on the allocation of output across its various uses - given a level of output, Y, what share goes towards final consumption versus paying the resource costs outlined in (17). If the amount of resources absorbed by non-consumption activities changes due to M&A, consumer outcomes may behave differently than industrial outcomes. Consider the identity $C = Y \times \frac{C}{Y}$. Taking logs and differencing gives the change in consumption as a function of the change in output and in consumption's share:

$$\Delta \log C = \Delta \log Y + \Delta \log \frac{C}{Y} \tag{20}$$

[35] If $\widehat{z}_{cf} > \widehat{z}$, an analogous, though slightly modified, expression holds.

5.1 An Economy Without M&A

To evaluate the effects of M&A, I compare the aggregate outcomes in the estimated economy to a counterfactual economy with no merger activity. Because the quantitative magnitudes are sensitive to the value of α (although the qualitative results are not), I report results for three values - a baseline value of 0.66, a low value of 0.5, and a high of 0.7 (expression (18) shows that α directly determines the elasticity of output with respect to changes in $\overline{Z}$ and M that arise from M&A). These span a broad range used in the literature. The baseline value of 0.66 matches an aggregate labor share of income of two-thirds.[36] The remaining counterfactuals work with this value.

Table 8 reports the changes in output and consumption when moving from the benchmark economy to one with no M&A.[37] The table shows (1) the potentially significant beneficial impact of M&A and (2) the important role of general equilibrium effects through changes in the entry/exit incentives of new entrepreneurs and the size distribution of firms. Eliminating M&A results in productivity (output) losses of between 11% and 26% and consumption losses of between 3% and 13%.[38] The largest portion of the productivity loss is due to changes in $\overline{Z}$, a reduction in average firm productivity. About 60% of this fall is on the intensive margin, i.e., comes from the loss of productivity gains created directly through merger, while 40% is due to reduced selection, i.e., M&A leads to increased selection of new entrants so that in its absence, lower quality entrepreneurs successfully enter, decreasing average productivity. The number of firms ("entry") can either increase or decrease, depending on the value of α, but the effects are generally modest (indeed, they are highest in the case of $\alpha = 0.5$, where M&A leads to more entry and so an even greater contribution to aggregate productivity). In addition to revealing a positive overall impact of M&A, an important conclusion is that indirect effects - i.e., changes in firm composition from selection and entry - compose a significant share of its total contribution.

The fall in consumption is considerably smaller than that in productivity/output. This is because M&A changes the allocation of output across its various uses. In particular, much of the output loss in the no-merger economy is offset by a rise in consumption's share of output. It turns out the wedge between output and consumption changes is not due to expenditures on search, which account for a negligible share of output (less than 1%). Rather, the existence of

[36] In a monopolistically competitive environment with CES preferences, the first two values correspond to an elasticity of substitution of 3 and 4, respectively. The sensitivity of aggregates to the curvature in production/demand is a recognized result in the firm dynamics literature, see, e.g., David et al. (2016).

[37] I outline the details of the computational algorithm used to compute counterfactuals in Online Appendix G.

[38] These values represent a comparison between two stationary equilibria not accounting for the transition path, and so should be interpreted as capturing the long-run effects of M&A.

Table 8: A No-Merger Economy

	$\alpha = 0.5$	$\alpha = 0.66$	$\alpha = 0.7$
Output/TFP	−26.4	−13.2	−11.3
Entry	−4.6	1.3	1.7
Intensive Margin	−13.5	−9.0	−8.1
Selection	−8.3	−5.6	−5.0
Consumption	−12.6	−3.7	−2.7
Consumption Share of Output	13.9	9.5	8.6

Notes: Table reports percentage changes in economic outcomes when moving from the benchmark economy to a no-merger economy.

the merger market generates a greater amount of unsuccessful attempted entry by entrepreneurs - i.e., "churn" - each of whom uses resources to pay entry costs that otherwise would have gone to households. The cost of churn ranges from 4% to 7% of output in the no-M&A economy and approximately doubles in the economy with M&A. Intuitively, the additional value stemming from their M&A prospects induces a large increase in the number of entrepreneurs attempting to enter the economy. Because M&A simultaneously makes successful entry more difficult through increased selection, it follows that churn must increase, i.e., more entrepreneurs pay the entry cost, draw an initial productivity below the threshold and exit immediately. This activity imposes large costs on the economy by absorbing a significant share of output. These results point to a key tradeoff between the productivity gains generated by M&A and the accompanying costs of devoting increased resources towards the churning of firms.[39]

The model gives a central role to M&A as a source of growth for acquiring firms by assuming that z is constant in between mergers. While clearly a stylized assumption, the data show that the idiosyncratic component of annual growth rates for the non-merging firms in my sample are, on average, close to zero (0.9%).[40] In contrast, the average one-year growth rate of acquiring firms is about 10%.[41] As pointed out by Haltiwanger et al. (2013) and Acemoglu et al. (2013), relatively large, mature firms tend to exhibit low rates of growth (indeed both papers report similarly low growth rates for these groups of firms). Further, Acemoglu et al. (2013) show that these firms exhibit low rates of R&D investment, one measure of firm-level investments in internal growth.[42] Because these are the types of firms in the Compustat sample - and,

[39]One example is the dot-com boom, which featured large rates of entry, anecdotal evidence suggests much of which was fueled by the potential to participate in the merger market, followed closely by large rates of exit. Though this may have ultimately led to gains in aggregate efficiency through successful entrants and the reallocation occurring via M&A, a significant amount of resources were absorbed by the activities of the ultimately unsuccessful entrepreneurs.

[40]To calculate this value, I regress (the log of) firm-level profitability on a full set of industry-year fixed effects and use the residual as the component that is idiosyncratic to the firm.

[41]Over a three year span, non-merging firms grow an average of 2.3%; acquiring firms 19%.

[42]There is a recent literature directly investigating the relationship between M&A and corporate innovation.

importantly, the ones that predominantly become acquirers - within-firm growth rates tend to be low and M&A seems to be a key channel for continued growth.

5.2 The Role of Sorting

Next, I explore the role of sorting in leading to merger gains, both at the firm and aggregate level. I perform two exercises, one in partial equilibrium and the second in general equilibrium. In the first, I compute the average combined firm-level gains across transactions in the benchmark economy as

$$\mathbb{E}\left[\frac{\Sigma\left(z_a, z_t\right)}{V\left(z_a\right)+V\left(z_t\right)}\right]=\int\int \Phi\left(\Sigma\left(z_a, z_t\right)\right)\left(\frac{\Sigma\left(z_a, z_t\right)}{V\left(z_a\right)+V\left(z_t\right)}\right)\Gamma\left(z_t\right)\Lambda\left(z_a\right)$$

I then perturb this equation in three ways - first, I eliminate sorting altogether and match firms completely randomly. To do so, I set the arrival rates λ and μ to exogenous constants (they then play no role in the equation) and additionally remove firms' option to reject unprofitable partners, i.e., I set $\Phi\left(\cdot\right)$ equal to one for all transactions. In other words, firms meet one another at a common rate and merge with any potential partner. Next, I examine each of these margins separately, i.e., I allow firms to reject potential partners but assume common arrival rates, and then I keep arrival rates fixed at their benchmark levels but remove the option to reject matches.[43] I report the results in Table 9. The top row shows the average combined merger gains under each of these scenarios, the second row the loss in percentage gains compared to the benchmark economy and the third row the share of the loss accounted for by each margin. Average combined gains are about 7% in the benchmark economy. Removing sorting altogether, i.e., random matching, leads to an average combined loss of about 10%, 17 percentage points lower than the benchmark. The endogenous selection of partners plays the dominant role in generating merger gains - forcing firms to merge with any potential match while holding arrival rates fixed at their benchmark levels leads to an average loss of about 8%, which accounts for almost 90% of the effect of sorting. The role of endogenous search intensities is smaller but still significant - with common arrival rates, but allowing for endogenous matching decisions, the combined gains fall from about 7% to 5.5%, accounting for about 10% of the total loss in the no-sorting scenario.

To quantify the effect of sorting on aggregate outcomes, I solve for the aggregates in a

The bulk of these studies (although not all) find that an active M&A market tends to lead to higher levels of firm innovation. For example, in a detailed study, Phillips and Zhdanov (2012) find that the possibility of an acquisition amplifies the potential gains from innovation, that firms' incentives to conduct R&D to innovate increase with industry acquisition activity and that R&D increases with the probability of becoming a target.

[43] Lentz (2010) points out the distinction between these two notions of sorting, i.e., through matching set variation or differing search intensities.

Table 9: The Role of Sorting - Firm-Level Merger Gains

	Benchmark	No Sorting	Constant Search	No Matching
Avg. Merger Gains	7.2	−10.3	5.5	−8.0
Loss vs Benchmark		−17.4	−1.7	−15.2
Share of Loss		100.0	9.5	87.0

Notes: Table reports average combined merger gains as a percentage of the sum of acquirer and target market values in (1) the benchmark economy, (2) with no sorting - constant search intensities and no endogenous match decision, (3) with constant search and endogenous matching and (4) endogenous search and no endogenous matching.

counterfactual scenario in which arrival rates are exogenous and common across firms. I set the average search intensity so that the total costs of search are the same as in the benchmark economy. Table 10 reports the results from this experiment. To gauge the magnitude of these effects, the table also reports the corresponding statistics when eliminating M&A altogether (from Table 8), and the proportion of the total that is attributable to endogenous search. The results show that firms' endogenous search decisions plays a key role in leading to aggregate gains from M&A. For example, total output losses from eliminating M&A completely are about 13%. Endogenous search accounts for about 60% of this value, about one-third of the direct productivity gains from merger and one-half of the changes in selection. Moreover, eliminating endogenous search causes firm entry to fall (whereas it rose slightly in the no-M&A economy under this parameterization), further lowering aggregate output. Strikingly, removing firms' endogenous search decisions leads to a fall in steady state consumption about 80% of that when eliminating M&A altogether.

I have also computed the counterfactual outcomes in the case of no sorting at all, i.e., with common exogenous arrival rates and no ability of firms to reject unprofitable partners. In this case, the changes in aggregates are even more extreme, for example, consumption falls by about 12% - four times the amount when completely eliminating M&A. In other words, removing the ability of firms to sort by choosing search intensities and selecting profitable partners is extremely costly - the economy would be better off with no M&A activity at all!

5.3 Efficiency and Policy

In Appendix D, I set up and derive the solution to the problem of a hypothetical social planner constrained by the same search frictions. Although the environments differ in several respects, many of the insights in Shimer and Smith (2001) go through here. In the presence of heterogeneity, the decentralized equilibrium is not efficient, meaning there is a scope for policy to improve economic outcomes. The planner's optimality conditions give a sharp characterization

Table 10: The Role of Sorting - Aggregate Outcomes

	Constant Search	No Mergers	Share
Output/TFP	-8.1	-13.2	61.0
Entry	-2.4	1.3	-
Intensive Margin	-3.1	-9.0	34.4
Selection	-2.7	-5.6	49.5
Consumption	-3.1	-3.7	83.5
Consumption Share of Output	5.0	9.5	52.3

Notes: Table reports percentage changes in economic outcomes when (1) eliminating endogenous search intensities, (2) when eliminating M&A and (3) the proportion of (2) accounted for by (1).

of the inefficiencies. In the case where $\theta_t < 1,\ \theta_a = 1$, the planner's conditions are[44]

$$\begin{aligned}(r+\delta) W(z) &= \pi(z) - C(\lambda(z)) - C(\mu(z)) \\ &+ \lambda(z)\theta_a \mathbb{E}[\Sigma(z, z_t)] + \mu(z)\theta_t \mathbb{E}[\Sigma(z_a, z)] - \mu(z)\theta_t \mathbb{E}[\Sigma(z_a, z_t)]\end{aligned} \tag{21}$$

where $\lambda(z)$ and $\mu(z)$ satisfy

$$\begin{aligned}C'(\lambda(z)) &= \theta_a \mathbb{E}[\Sigma(z, z_t)] \\ C'(\mu(z)) &= \theta_t \mathbb{E}[\Sigma(z_a, z)] - \theta_t \mathbb{E}[\Sigma(z_a, z_t)]\end{aligned} \tag{22}$$

Here, $W(z)$ denotes the social marginal value of a firm of type z, i.e., the value to the economy of a marginal increase in the mass of firms of that type; $\Sigma(z_a, z_t) = W(z_m) - W(z_a) - W(z_t)$ the social surplus from a match; $\mathbb{E}[\Sigma(z, z_t)] = \int \max\{\Sigma(z, z_t), 0\}\Gamma(z_t)$ the expected social surplus from matching firm z as an acquirer (the expectation is with respect to the target the firm might meet, z_t) and similarly, $\mathbb{E}[\Sigma(z_a, z)]$ the expected surplus from matching firm z as a target; and $\mathbb{E}[\Sigma(z_a, z_t)] = \int\int \max\{\Sigma(z_a, z_t), 0\}\Gamma(z_t)\Lambda(z_a)$ the expected surplus from a match across all types, i.e., the expectation is unconditional over all z_a and z_t. The planner also solves analogous equations to the free entry and entry threshold conditions

$$\begin{aligned}\int W(z)\,dF(z) &= c_e \\ W(\hat{z}) &= 0\end{aligned} \tag{23}$$

Compare the planner's conditions to the equilibrium conditions (8), (9), (13) and (14). The social marginal value of a firm has three components: first is the static contribution to

[44] This is the case in the benchmark equilibrium, where $\theta_t = 0.86$. Appendix D derives analogous expressions in the cases where $\theta_a < 1,\ \theta_t = 1$ and $\theta_a = \theta_t = 1$.

aggregate production less the costs of search. These are the same terms as enter the private value.[45] Second, the firm may merge, either as an acquirer or target, generating expected social surplus $\mathbb{E}\left[\Sigma\left(z, z_t\right)\right]$ or $\mathbb{E}\left[\Sigma\left(z_a, z\right)\right]$, respectively. However, the social value of the firm depends on the entire surplus, $\Sigma\left(\cdot\right)$, whereas the private value depends only on the firm's share, $\beta\Sigma\left(\cdot\right)$ or $\left(1-\beta\right)\Sigma\left(\cdot\right)$. This is a "thick markets" externality - firm private values depend only on the share of the match surplus the firm captures, whereas firm social values include the full match surplus. The third (negative) term only appears in the planner's value function, not in the private one and reveals a "congestion" externality - as the firm searches more intensively, it makes it more difficult for other firms to match. Congestion shows up on the side of the market where the matching rate is constrained by market tightness (in this example, I have assumed this is the target side). The social cost of this externality depends on the unconditional expected surplus across all matches.

These same effects show up in the optimal search intensities - the planner makes search a function of the entire surplus, internalizing the thick markets externality, and takes into account the effects of each firm's search on the matching rates of all other firms, internalizing congestion. The remainder of the optimality conditions are analogous to the equilibrium conditions, but the planner uses social rather than private values - the planner will form matches with non-negative social surplus, i.e., $W\left(z_m\right) - W\left(z_a\right) - W\left(z_t\right) \geq 0$, and chooses the extent and quality of new entrants to satisfy a type of free-entry condition - expected social values must be equal to the cost of entry - and entrant threshold condition - the planner has firms enter that have non-negative social value.

Appendix D derives a system of linear type-dependent taxes/subsidies on search intensities that decentralizes the social optimum. Denoting these taxes τ^{λ} and τ^{μ} and continuing to assume $\theta_t < 1, \theta_a = 1$, they take the form

$$\begin{aligned}
\tau^{\lambda} &= -\left(1-\beta\right)\theta_a\mathbb{E}\left[\Sigma\left(z, z_t\right)\right] \\
\tau^{\mu} &= -\beta\theta_t\mathbb{E}\left[\Sigma\left(z_a, z\right)\right] + \theta_t\mathbb{E}\left[\Sigma\left(z_a, z_t\right)\right]]
\end{aligned}$$

The planner would like to subsidize search to address the thick markets externality and tax search to address congestion (search on the short side of the market is always subsidized, since there are no congestion effects). The taxes will be zero only if the distribution of surplus across matches is degenerate, which is not generally the case. Firms that have high expected social surplus of matching will tend to be subsidized and firms that have low expected social surplus will tend to be taxed.

[45] Appendix D proves that the firm's contribution to aggregate output is exactly the constant that enters the equilibrium profit function.

I illustrate these effects using some simple policies. First, I investigate a system of flat taxes on merger gains, i.e., the merging firms only receive (and split) a fraction $1-\tau$ of the gains (I assume tax revenues are rebated to the household in lump-sum form). In the first column of Table 11, I consider a tax rate of 40%. Simulations show that this is approximately the rate that maximizes the positive effects of this type of policy - steady state consumption increases almost 3%.[46] In this case, productivity actually falls - although intensive margin gains increase, the tax reduces entry and selection, lowering entrant churn and shifting resources to households. The rate of merger falls by about 30% (from 3.7% to 2.6%). Next, I consider the opposing policy - a subsidy valued at 20% of merger gains. Here, the results are flipped - consumption fall about 2.5%. The rate of merger increases, but so does entrant churn, diverting resources from the household. This exercise provides one reference point for the potential gains from policy and highlights that although M&A has an overall positive impact, it is not necessarily the case that "more is better" - with only this policy instrument, the planner would prefer to reduce the merger rate to limit entrant churn and reallocate the saved resources to households.

Table 11: The Effects of Policy

	Tax Rate		Prob (Blocked)	
	40%	-20%	12%	80%
Output/TFP	−4.2	3.9	−1.5	−9.1
Entry	−3.9	2.4	−1.0	−1.5
Intensive Margin	1.8	−0.1	0.0	−3.5
Selection	−2.1	1.6	−0.5	−4.2
Consumption	2.7	−2.4	1.0	−1.2
Consumption Share of Output	6.9	−6.3	2.5	7.9
Merger Rate	−32.4	28.2	−16.0	−62.1

Notes: Table reports percentage changes in economic outcomes when introducing a system of flat taxes/subsides and under size-dependent policies.

In the last two columns of Table 11, I consider a different type of policy - namely, size-dependent policies under which proposed transactions are subject to a probability of being blocked by government action, in the spirit of antitrust enforcement. The tradeoff between antitrust concerns and efficiency gains is a key consideration in merger policy, but it has typically been difficult to measure the foregone efficiency gains - doing so requires a theory of merger gains.[47] Further, to quantify the long-run changes in aggregate outcomes requires an

[46] Further, this is approximately the statutory corporate tax rate in the US (inclusive of state and local taxes), suggesting it is not an unreasonable level. See, for example, https://stats.oecd.org/index.aspx?DataSetCode=Table_II1.

[47] See, for example, the FTC and DOJ Horizontal Merger Guidelines, available at https://www.ftc.gov/sites/default/files/attachments/merger-review/100819hmg.pdf and the in-depth dis-

environment with endogenous merger formation and entry, features that are new to the framework here. Although the model does not feature antitrust concerns, we can use it to calculate an estimate of the foregone efficiency gains due to these policies, which represents a potential upper bound on their detrimental impact.

I investigate two policies - in the first, transactions exceeding \$1 billion are subject to a 12% probability of being. This value is broadly motivated by current enforcement.[48] The third column of Table 11 suggests the effects of this policy are modest and can actually have a small positive impact, even without antitrust concerns - steady state consumption increases by 1% and the detrimental effect on productivity, which falls 1.5%, is limited. Intuitively, mergers large enough to warrant antitrust concerns, i.e., above \$1 billion in value, are relatively rare (6% of mergers in the model and about 4.5% in the data). Preventing a small fraction of them has little effects on aggregate efficiency. Further, due to the inherent inefficiencies in the M&A market, the policy can actually work to improve aggregate outcomes even absent antitrust concerns. The last column of Table 11 shows that an almost prohibitive policy can have more substantial consequences. Blocking 80% of mergers valued over \$1 billion leads to losses in consumption and output of 1% and 9%, respectively.[49] The largest individual component of these changes comes through reduced selection - by preventing the growth of large firms, the policy enables less productive entrepreneurs to successfully enter the economy, reducing aggregate productivity.[50] These findings provide some insight into the tradeoff between antitrust enforcement and efficiency gains faced by the policymaker - at current enforcement levels, the foregone efficiency gains seem to be modest; the losses can be more substantial under under relatively stringent policies, where a large portion is due to changes in the characteristics of new entrepreneurs.

5.4 Merger Waves

A well documented phenomenon surrounding M&A is the existence of merger "waves," i.e., high frequency periods of increased merger activity.[51] In this section, I use the model to investigate three possible candidates for merger booms - increases in "merger efficiency," A, falling search

cussion in Farrell and Shapiro (2001).

[48]For example, over the past 10 years, an average of 11.5% of transactions exceeding \$1 billion were subject to further investigation by the antitrust authorities through a "second request" for further information (the corresponding figure for transactions below \$1 billion is only about 2.5%). See, e.g., Hart-Scott-Rodino Annual Reports available at http://www.ftc.gov/bc/anncompreports.shtm.

[49]The merger rate falls substantially as well, mostly because the greater presence of low productivity firms leads to increased congestion among small firms, who greatly reduce their M&A activities.

[50]Relatedly, Hsieh and Klenow (2014) show that distortions hindering firm growth can have significant negative effects on the selection of new entrants.

[51]For example, Jovanovic and Rousseau (2008) and Harford (2005) study merger waves and Eisfeldt and Rampini (2006) the procyclical properties of M&A.

costs, B, and shocks to the discount rate, r.[52] I compute elasticities of merger market variables with respect to each of these parameters by solving for the equilibrium outcomes under small perturbations to the parameter and calculating $\frac{d\log(x)}{d\log(y)}$, where the numerator is the percentage change in the outcome and the denominator the percentage change in the driving variable.

Table 12 reports the results. Beginning with the middle column, the table shows that merger activity is not sensitive to small changes in search costs. The first order conditions on search show that holding expected gains constant, the elasticity of search with respect to the scale of costs is $\frac{1}{\eta-1}$. Because the estimate of η is rather large (i.e., search costs seem to be quite convex), small changes in these costs give firms little incentive to adjust their behavior. Additional simulations show that reductions in search costs do lead to increases in M&A, but that these changes need to be substantial. These findings suggest that unless search costs exhibit large variation at high frequencies, they may not be the most likely candidate behind merger waves (although cost reductions might be a more important factor behind the secular increase in M&A discussed in footnote 1, due, for example, to advances in information technology, etc.).

In contrast, the first column of Table 12 shows that changes in merger efficiency, A, have more significant effects. The elasticity of the merger rate is 0.45 and of the raw count of mergers about 1.8, suggesting that the volume of M&A activity is sensitive to changes in A. Further, entrepreneur entry also shows a high elasticity with respect to merger efficiency, suggesting that variation in this term can reconcile simultaneous jumps in merger activity and entry. Why do shocks to merger efficiency lead to large responses while cost shocks do not? Intuitively, changes in costs only affect firms' actions on the merger market, which is generally a modest part of overall firm value. Greater productivity gains from merger lead to permanently higher levels of profitability and so garner a larger reaction in pursuit of those gains. A fall in the discount rate leads to qualitatively similar effects, although the quantitative magnitudes are smaller. However, discount variation is known to be large, even at high frequencies, which would lead to larger effects on M&A.[53] In sum, the results point to changes in merger efficiency and discount rates as two candidates behind merger waves. A more in-depth exploration of these forces (and perhaps others not examined here) could be a useful application of the framework.

6 Conclusion

This paper developed a theory of firm-level merger activity and links M&A to aggregate economic outcomes through changes in selection, entrepreneurship and the size distribution of

[52] Times of high M&A activity tend to be times of high market valuations, see, e.g., Jovanovic and Rousseau (2001) and Maksimovic and Phillips (2001).

[53] See, e.g., Cochrane (2011).

Table 12: Elasticities of Mergers and Entry

	$\frac{\Delta \log x}{\Delta \log A}$	$-\frac{\Delta \log x}{\Delta \log B}$	$-\frac{\Delta \log x}{\Delta \log r}$
Number of Mergers	1.763	0.005	0.280
Merger rate	0.452	0.003	0.118
Entry	1.311	0.002	0.162

Notes: Table reports elasticities of merger outcomes and entry with respect to (1) merger efficiency, (2) search costs and (3) discount rates.

firms. The theory is consistent with a rich set of micro-level facts on US M&A including sorting among merging firms, a substantial merger premium and serial acquisition. It uses these facts to shed new light on the nature of merger gains, which come through both synergies, i.e., complementarities between merging firms, and productivity improvements of target firms. The results point to a significant beneficial impact of M&A on aggregate outcomes, much of which stem from the endogenous sorting of merging firms. Because of inefficiencies in the search and matching process, the findings suggest that even simple policies can have significant positive effects on economic outcomes.

There are several promising directions for future research. In the theory, the process of merging is captured by a parsimonious technology by which firms combine. My analysis relates the shape of the technology to the nature of merger gains and empirical patterns in M&A. Deriving explicit microfoundations for this process would shed further light on the underlying mechanisms through which firm combine, the boundaries of the firm and, more broadly, may be useful to guide future work investigating knowledge diffusion across firms. Recent empirical evidence suggests a rich set of interactions between M&A and internal avenues for innovation and growth, indeed, that an active M&A market tends to encourage R&D and innovation. Quantifying these relationships within a rich setting of endogenous growth and M&A, empirically disciplined using micro-data on firm-level innovation, would be a valuable next step. My work has touched on issues of efficiency and time variation in M&A. Using the framework to further explore the effects of merger policy may lead to new considerations regarding interventions in the M&A market. Investigating the high-frequency behavior of M&A may generate new insights into merger waves and the procyclicality of merger activity.

References

Acemoglu, D., U. Akcigit, N. Bloom, and W. R. Kerr (2013): "Innovation, Reallocation and Growth," Working Paper 18993, National Bureau of Economic Research.

Ahern, K. (2012): "Bargaining Power and Industry Dependence in Mergers," *Journal of Fi-*

nancial Economics, 103, 530–550.

Andrade, G., M. Mitchell, and E. Stafford (2001): "New Evidence And Perspectives on Mergers," *Journal of Economic Perspectives*, 15, 103–120.

Atakan, A. E. (2006): "Assortative Matching with Explicit Search Costs," *Econometrica*, 74, 667–680.

Atkeson, A. and A. Burstein (2010): "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118, 433–484.

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293, 1818–1820.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93, 1268–1290.

Betton, S., B. Eckbo, and K. Thorburn (2008): "Corporate Takeovers," *Handbook of Corporate Finance: Empirical Corporate Finance*, 2, 289–427.

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.

Boone, A. and J. Mulherin (2007): "How Are Firms Sold?" *Journal of Finance*, 62, 847–875.

Braguinsky, S. and B. Jovanovic (2004): "Bidder Discounts and Takeover Premia," *American Economic Review*, 94, 46–56.

Braguinsky, S., A. Ohyama, T. Okazaki, and C. Syverson (2015): "Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry," *The American Economic Review*, 105, 2086–2119.

Chade, H., J. Eeckhout, and L. Smith (2017): "Sorting Through Search and Matching Models in Economics," *Journal of Economic Literature*, 55, 493–544.

Cochrane, J. H. (2011): "Presidential Address: Discount Rates," *The Journal of Finance*, 66, 1047–1108.

David, J. M., H. A. Hopenhayn, and V. Venkateswaran (2016): "Information, Misallocation, and Aggregate Productivity," *The Quarterly Journal of Economics*, 131, 943–1005.

DePamphilis, D. (2009): *Mergers, Acquisitions, and Other Restructuring Activities: An Integrated Approach to Process, Tools, Cases, and Solutions*, London, U.K.: Academic Press.

Eckbo, B. E. (2014): "Corporate Takeovers and Economic Efficiency," *Annual Review of Financial Economics*, 6, 51–74.

Eeckhout, J. and P. Kircher (2010): "Sorting and Decentralized Price Competition," *Econometrica*, 78, 539–574.

——— (2011): "Identifying Sorting—In Theory," *The Review of Economic Studies*, 78, 872–906.

Eisfeldt, A. and A. Rampini (2006): "Capital Reallocation and Liquidity," *Journal of Monetary Economics*, 53, 369–399.

——— (2008): "Managerial Incentives, Capital Reallocation, and the Business Cycle," *Journal of Financial Economics*, 87, 177–199.

Faria, A. (2008): "Mergers and the Market for Organization Capital," *Journal of Economic Theory*, 138, 71–100.

Farrell, J. and C. Shapiro (2001): "Scale Economies and Synergies in Horizontal Merger Analysis," *Antitrust Law Journal*, 68, 685–710.

Hagedorn, M., T. H. Law, and I. Manovskii (2017): "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85, 29–65.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013): "Who Creates Jobs? Small Versus Large Versus Young," *Review of Economics and Statistics*, 95, 347–361.

Harford, J. (2005): "What Drives Merger Waves?" *Journal of Financial Economics*, 77, 529–560.

Hopenhayn, H. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150.

Hsieh, C.-T. and P. J. Klenow (2014): "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, 129, 1035–1084.

Jovanovic, B. and P. Rousseau (2002): "The Q-Theory of Mergers," *American Economic Review*, 92, 198–204.

——— (2008): "Mergers as Reallocation," *Review of Economics and Statistics*, 90, 765–776.

Jovanovic, B. and P. L. Rousseau (2001): "Mergers and Technological Change: 1885-1998," *mimeo*, Vanderbilt University Department of Economics.

Lentz, R. (2010): "Sorting by Search Intensity," *Journal of Economic Theory*, 145, 1436–1452.

Li, K. and N. R. Prabhala (2007): "Self-Selection Models in Corporate Finance," *Handbook of Corporate Finance: Empirical Corporate Finance*, 3.

Li, X. (2013): "Productivity, Restructuring, and the Gains from Takeovers," *Journal of Financial Economics*, 109, 250–271.

Lucas, R. (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9, 508–523.

Lucas, R. E. and B. Moll (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122, 1–51.

Maksimovic, V. and G. Phillips (2001): "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?" *Journal of Finance*, 56, 2019–2065.

Maksimovic, V., G. Phillips, and N. R. Prabhala (2011): "Post-Merger Restructuring and the Boundaries of the Firm," *Journal of Financial Economics*, 102, 317–343.

Malmendier, U., M. M. Opp, and F. Saidi (2016): "Target Revaluation After Failed Takeover Attempts: Cash Versus Stock," *Journal of Financial Economics*, 119, 92–106.

Martos-Vila, M. (2008): "The Search for Corporate Control," *mimeo*, UCLA.

Masulis, R., P. Swan, and B. Tobiansky (2008): "Do Wealth Creating Mergers and Acquisitions Really Hurt Acquirer Shareholders?" *mimeo*, School of Banking and Finance, Australian School of Business, UNSW.

Melitz, M. (2003): "The Impact Of Trade On Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

Moeller, S., F. Schlingemann, and R. Stulz (2004): "Firm Size and the Gains from Acquisitions," *Journal of Financial Economics*, 73, 201–228.

Officer, M. S. (2003): "Termination Fees in Mergers and Acquisitions," *Journal of Financial Economics*, 69, 431–467.

Perla, J. and C. Tonetti (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, 52–76.

Phillips, G. M. and A. Zhdanov (2012): "R&D and the Incentives from Merger and Acquisition Activity," *The Review of Financial Studies*, 26, 34–78.

Rhodes-Kropf, M. and D. Robinson (2008): "The Market for Mergers and the Boundaries of the Firm," *Journal of Finance*, 63, 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, 561–603.

Schoar, A. (2002): "Effects of Corporate Diversification on Productivity," *Journal of Finance*, 57, 2379–2403.

Shimer, R. and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369.

——— (2001): "Matching, Search, and Heterogeneity," *BE Journal of Macroeconomics*, 1, 5.

Wang, W. (2012): "Are Takeovers Really Bad Deals for the Acquirers?" *mimeo*, Wisconsin School of Business.

Welch, I. (2009): *Corporate Finance: An Introduction*, New York: Prentice Hall.

Appendix

A Data

I select from SDC all domestic transactions announced between 1977 and 2009 with a nominal deal value of at least \$1 million. I include only completed transactions; those not classified as hostile (only about 300 transactions are classified as hostile takeovers); those in which the acquirer newly gains majority control of the target, i.e., the acquirer must own less than 50% of the target prior to the merger and over 50% after; and those with relevant ownership status (excluding, for example, government-owned entities). After this process, and eliminating several observations with obvious data entry errors, there are 57,858 transactions. For each transaction, I obtain the following data: transaction value (total value of consideration paid by the acquirer, excluding fees and expenses), premium (premium of offer price to target closing stock price 4 weeks prior to the original announcement date), and pre-merger performance variables including net sales, employment, PP&E, EBITDA, and market value. Data availability differs across the

SDC variables. In Table 13, I show the number of transactions with available data for acquirers, targets, and both, for each dimension of analysis.

Table 13: SDC Data Availability

	Acquirer	Target	Both
Sales	31,736	18,541	12,251
Employment	28,050	6,138	3,957
PP&E	28,792	10,095	6,672
EBITDA	26,424	8,208	5,080
Market Value	25,38	6,969	4,112
Premium	*	*	6,474

Turning to Compustat, I obtain data on the universe of firms contained in the CRSP/Compustat merged database (CCM) from 1977 to 2009. This yields a total of 210,275 observations. The SDC to Compustat match is not straightforward since the two databases use different company identifiers. The most specific identifier provided by SDC is the 6-digit CUSIP for both parties in each transaction. This is not sufficient for the match, however, because Compustat only records the most recent CUSIP rather than a CUSIP history. Because of this, matching on CUSIP may result in missed pairs and erroneous matches.

To perform the match, I use the CRSP translator to associate 6-digit CUSIPs from SDC with the CRSP company identifier. I then match this identifier with the CCM database, which already associates the CRSP identifier with the set of Compustat firms. I follow this process for both acquirers and targets. I associate transactions with the Compustat data for the fiscal year preceding the year of merger announcement. The set of successful matches corresponds closely to the set of firms classified as public in SDC. I obtain data on net sales, employees, PP&E (net of depreciation), EBITDA, and market value, where I calculate the latter as the product of common shares outstanding and the closing price at fiscal year end. Table 14 shows availability of the Compustat data.

Table 14: Compustat Data Availability

	All Firms	Acquirers	Targets	Both
Sales	191,992	30,453	6,828	4,465
Employment	179,787	28,597	6,186	3,862
PP&E	187,270	28,368	6,608	4,284
EBITDA	152,671	25,472	5,525	3,478
Market Value	206,309	30,890	6,913	4,548

Macroeconomic data are obtained from standard sources. US GDP and stock of fixed assets are from the Bureau of Economic Analysis (http://www.bea.gov/). The CPI is from the Bureau of Labor Statistics (http://www.bls.gov/).

B Merger Gains and Merger Patterns

B.1 Proofs of Propositions

Proposition 1. Conjecture that the gains from merger are constant and strictly positive across all firm types, i.e., $\Sigma(z_a, z_t) = \bar{\Sigma} > 0 \ \forall z_a, z_t$. Then all meetings will result in a completed transaction and we can write the value function as

$$rV(z) = \pi(z) - C(\lambda(z)) - C(\mu(z)) + \lambda(z)\theta_a\beta\bar{\Sigma} + \mu(z)\theta_t(1-\beta)\bar{\Sigma}$$

The first order conditions governing optimal search in (9) give

$$C'(\lambda(z)) = \theta_a\beta\bar{\Sigma}, \quad C'(\mu(z)) = \theta_t(1-\beta)\bar{\Sigma}$$

which shows that the choice of λ and μ are common across firms and independent of z. Denoting these common search intensities as $\bar{\lambda}$ and $\bar{\mu}$, we can rewrite the value function as

$$rV(z) = \pi(z) - C\left(\bar{\lambda}\right) - C(\bar{\mu}) + \bar{\lambda}\theta_a\beta\bar{\Sigma} + \bar{\mu}\theta_t(1-\beta)\bar{\Sigma} \tag{24}$$

Recall that the gains from merger between a type z_a acquirer and type z_t target are $\Sigma(z_a, z_t) = V(z_m) - V(z_a) - V(z_t)$. Using (24), it is straightforward to show that merger gains equal

$$\frac{\pi(z_m) - \pi(z_a) - \pi(z_t) - \bar{\lambda}\theta_a\beta\bar{\Sigma} - \bar{\mu}\theta_t(1-\beta)\bar{\Sigma} + C\left(\bar{\lambda}\right) + C(\bar{\mu})}{r}$$

Under the assumption that the merger technology displays no gains from bundling, we have

$$\pi(z_m) - \pi(z_a) - \pi(z_t) = c_f$$

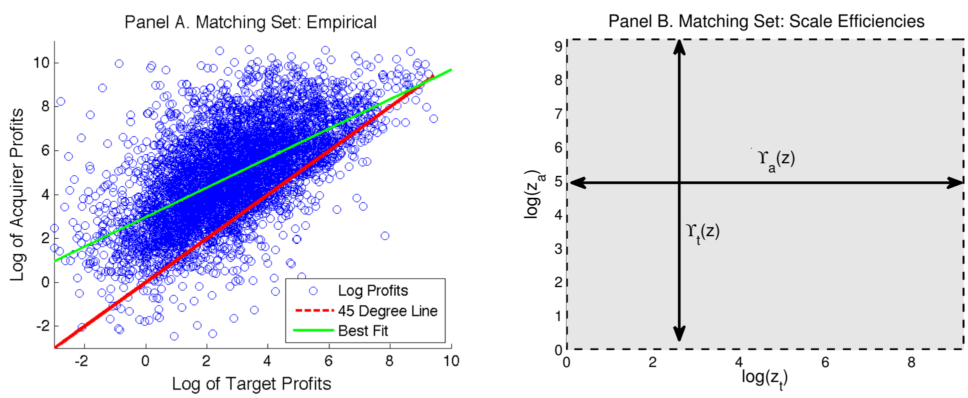
that is, the only gain in flow profits from merging is a single fixed cost savings. Then,

$$\Sigma(z_a, z_t) = \frac{c_f - \bar{\lambda}\theta_a\beta\bar{\Sigma} - \bar{\mu}\theta_t(1-\beta)\bar{\Sigma} + C\left(\bar{\lambda}\right) + C(\bar{\mu})}{r} = \bar{\Sigma} > 0$$

Thus, we have proved our initial conjecture that the gains from merger are positive and constant across all firm types such that every meeting will result in merger.

Because each firm searches with the same intensities, and the effective meeting rates on the two sides of the market must equate, each firm has an equal probability of meeting a particular partner as an acquirer or a target. That is, the rate at which acquirer z_1 meets target z_2 equals the rate at which acquirer z_2 meets target z_1. That meetings are random and all result in a completed transaction, that all firms choose the same search intensities, and that each transaction is reflected by the opposite transaction with the roles reversed in equal weight together imply that (i) the mean and median differences between acquirers and targets are zero, (ii) the correlations between acquirers and targets are zero, and (iii) the median acquirer and median target are the same as the median firm. $\square$

Figure 6 compares the joint distribution over the profits of acquirers and targets from the data to the predictions of this theory. In Panel A, I replicate the empirical distribution presented in Figure 1, i.e., the empirical matching set. In Panel B, I show the matching set that results from a pure scale efficiency theory. The horizontal arrows represents the acceptance region for acquirers, that is, an acquirer is willing to merge with any firm along the length of the arrow. The vertical arrows analogously shows the acceptance region for targets. The matching set is shaded. The figure shows that the entire domain becomes the matching set, that is, all firms are willing to merge with all others. Because search intensities are equal across firms, the entire domain will be populated by a random array of mergers. Clearly, this technology predicts merger patterns (or lack thereof) that are quite far from the empirical ones.



Notes: Panel A displays the empirical matching set between acquirers and targets. Panel B displays the matching set under a theory of scale efficiencies. The shaded area denotes the matching set, which is the entire domain. Horizontal segments represent the acceptance regions of acquirers, $\Upsilon_a(z)$, and vertical segments the acceptance region of targets, $\Upsilon_t(z)$.

Figure 6: A Theory of Scale Efficiencies

Proposition 2. The symmetry of the merger technology along with the definition of the combined gains in (3) immediately imply that $\Sigma(z_1, z_2) = \Sigma(z_2, z_1)$ and that matching sets are symmetric around the 45° line. Conjecture that $\lambda(z) = K\mu(z), K > 1$, that is, for each firm, search intensity on the acquiring side of the market is some constant proportion of search intensity on the the target side. Then, $\int \lambda(z)\, dG(z) = K \int \mu(z)\, dG(z)$, that is, the aggregate search intensity of acquirers is the same proportion of the aggregate search intensity of targets. This implies $\theta_a < 1$ and $\theta_t = 1$. Expected gains for a type z target conditional on meeting a prospective buyer are

$$\begin{aligned}
E[\Sigma_t(z_a, z)] &= (1-\beta)\int \max\{\Sigma(z_a, z), 0\}\, \Lambda(z_a) \\
&= (1-\beta)\int \max\{\Sigma(z_a, z), 0\}\, \frac{\lambda(z_a)\, dG(z_a)}{\int \lambda(z)\, dG(z)} \\
&= (1-\beta)\int \max\{\Sigma(z, z_a), 0\}\, \frac{K\mu(z_a)\, dG(z_a)}{K\int \mu(z)\, dG(z)} \\
&= (1-\beta)\int \max\{\Sigma(z, z_a), 0\}\, \frac{\mu(z_a)\, dG(z_a)}{\int \mu(z)\, dG(z)} \\
&= (1-\beta)\int \max\{\Sigma(z, z_a), 0\}\, \Gamma(z_a) \\
&= \frac{1-\beta}{\beta} E[\Sigma_a(z, z_a)]
\end{aligned}$$

Expected gains conditional on meeting a prospective buyer are a constant multiple of the expected gains conditional on meeting a prospective target, and simply depends on the ratio of bargaining powers. Note that in the third line, I have used the symmetry assumption on the technology, as well as the initial conjecture that search intensities are in constant proportion. From the first order conditions, (9), and the homogeneity of the cost function, if expected gains are in constant proportion, than search intensities are as well, verifying the initial conjecture. Similar reasoning holds for the cases of $K < 1$ and $K = 1$.

From (7), the rates at which a type z_1 acquirer meets a type z_2 target, and in reverse, a type z_1 target meets a type z_2 acquirer are equal to

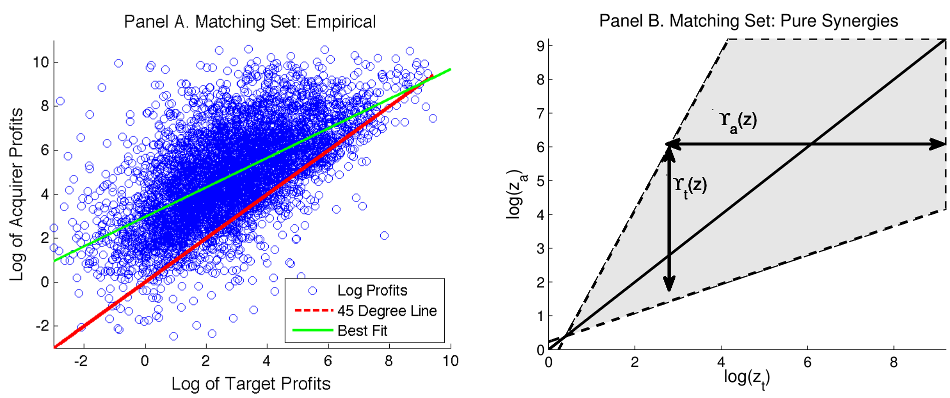
$$\lambda(z_1)\, \theta_a \frac{\mu(z_2)\, dG(z_2)}{\int \mu(z)\, dG(z)}, \quad \mu(z_1)\, \theta_t \frac{\lambda(z_2)\, dG(z_2)}{\int \lambda(z)\, dG(z)}$$

Substituting $\lambda(z) = K\mu(z)$ in both expressions, we obtain

$$K\mu(z_1)\, \theta_a \frac{\mu(z_2)\, dG(z_2)}{\int \mu(z)\, dG(z)}, \quad \mu(z_1)\, \theta_t \frac{\mu(z_2)\, dG(z_2)}{\int \mu(z)\, dG(z)}$$

which are equivalent since $\theta_t = 1$ and $K\theta_a = 1$. Thus, each transaction is reflected in equal weight by its counterpoint transaction with the roles reversed. It is then immediate that the mean and median differences between acquirers and targets are zero. □

Figure 7 compares the implied matching set to the empirical. For illustration purposes, I have assumed the conditions laid out in Section 3.2.3 hold, so the matching set will exhibit sorting. Under these conditions, the matching set more closely resembles the empirical one. However, in the absence of any asymmetry in the technology generating merger gains, the matching set is symmetric around the 45° line as will be the intensity of matches. In contrast, the empirical matching set is centered 2 log points, a factor of over 7, above the 45° line, with over 90% of transactions lying above the line. To capture the significant and pervasive differences between acquirers and targets, the theory must exhibit some asymmetry in the roles of acquirer and target in the technology creating gains from merger.

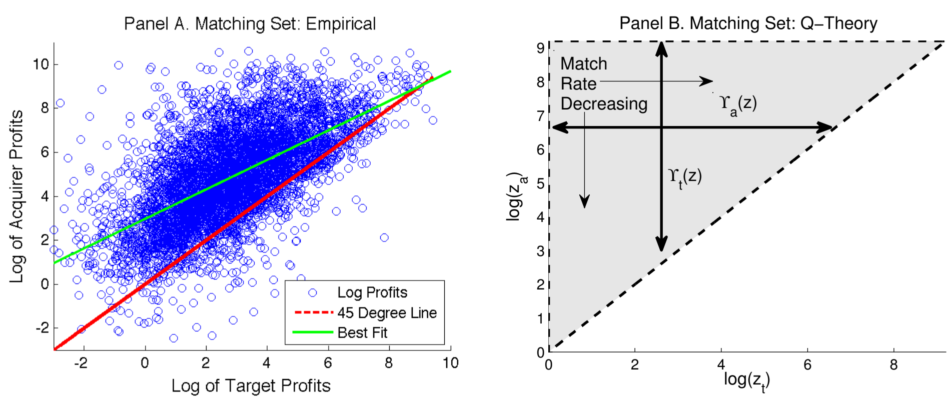


Notes: Panel A displays the empirical matching set between acquirers and targets. Panel B displays the matching set under a theory of purely synergistic mergers. The shaded area denotes the matching set. The dashed lines are the lower and upper bounds of the matching set and the solid line is the 45° line. Horizontal segments represent the acceptance regions of acquirers, $\Upsilon_a(z)$, and vertical segments the acceptance region of targets, $\Upsilon_t(z)$.

Figure 7: A Theory of Synergistic Mergers

Proposition 3. Assume that merger gains $\Sigma(z_a, z_t)$ are increasing in the difference between the acquirer and target $z_a - z_t$. Then for a given z_a, gains are decreasing in z_t, and so the set of acceptable targets $\Upsilon_a(z_a)$ is characterized by an upper threshold z_t^* such that $\Sigma(z_a, z_t^*) = 0$. That is, acquirer z_a will be willing to purchase any targets with $z \leq z_t^*$. It is straightforward to establish an analogous result for targets, that is, $\Upsilon_t(z_t)$ is characterized by a lower threshold z_a^* such that target z_t will sell itself to any acquirer with $z \geq z_a^*$. Together these imply low z targets and high z acquirers are in a greater share of matching sets. That gains are decreasing in z_t and increasing in z_a implies that expected gains conditional on meeting a candidate purchaser or target are also respectively decreasing and increasing in z. From the first order conditions governing optimal search (9), we see that search intensities $\mu(z)$ and $\lambda(z)$ must be decreasing and increasing in z respectively, that is, low z targets and high z acquirers search most intensively for partners. The fact that low z targets and high z acquirers are in a greater share of matching sets and search most intensively together imply that the rate at which firms are acquired $\mu(z)\,\theta_t \int \Phi(\Sigma_t(z_a, z))\,\Lambda(z_a)$ is decreasing in z and similarly the rate at which they make acquisitions $\lambda(z)\,\theta_a \int \Phi(\Sigma_a(z, z_t))\,\Gamma(z_t)$ is increasing in z. It is then immediate that (i) low z firms are overrepresented in the set of targets and high z firms in the set of acquirers and that (ii) the median target is below the median firm and the median acquirer above. Finally, the greatest rate of meeting take place between the highest z acquirer and the lowest z target, which is an acceptable match, giving that (iii) the highest rate of transaction occurs between low z targets and high z acquirers. □

Figure 8 illustrates the implied matching set and compares it to the empirical one. Any meeting of firms above the 45° line, i.e., where z_a is greater than z_t, results in a consummated merger, capturing the pervasive differences between acquirers and targets. It also captures to some degree positive sorting, since higher z acquirers are willing to take on higher z targets. However, the proposition shows that high z acquirers and low z targets transact at the most rapid rate, implying that the greatest density of matches should occur in the northwest corner of the domain, and that the density should be decreasing to the south and east from this corner. Turning to the empirical matching set, this match almost never occurs in the data. Moreover, the greatest density of matches is much more centered and decreases in the opposite direction, towards the northwest and southeast corners of the domain.



Notes: Panel A displays the empirical matching set between acquirers and targets. Panel B displays the matching set under a q-theory of mergers. The shaded area denotes the matching set. The dashed lines are the lower and upper bounds of the matching set, which are the 45° line and the highest z_a, respectively. Horizontal segments represent the acceptance regions of acquirers, $\Upsilon_a(z)$, and vertical segments the acceptance region of targets, $\Upsilon_t(z)$.

Figure 8: A Q-Theory of Mergers

B.2 Sorting

The setup is a variant of Eeckhout and Kircher (2011). In the absence of frictions, the period 2 matching market is competitive. The merger technology takes the symmetric form $s\left(z_a, z_t\right) = \left(z_a z_t\right)^{\varphi}$. Define the assignment of targets z_t to acquirers z_a as $\psi\left(z_t\right) = z_a$, meaning that target z_t matches with acquirer $z_a = \psi\left(z_t\right)$. Firms can reject the match, in which case they produce and earn profits z_a or z_t. Upon a match, the target firm gets payout $P\left(z_t\right)$, so that gains to the target are $\Sigma_t\left(z_t\right) = P\left(z_t\right) - z_t$ and gains to the acquirer are $\Sigma_a\left(z_a, z_t\right) = s\left(z_a, z_t\right) - z_a - P\left(z_t\right)$. An equilibrium in the period 2 matching market is a pair of functions $\left(\psi, P\right)$ such that

$$\Sigma_t\left(z_t, \psi\left(z_t\right)\right) + \Sigma_a\left(\psi^{-1}\left(z_a\right), z_a\right) \geq s\left(z_a, z_t\right) - z_a - z_t \;\forall\; z_a, z_t$$

i.e., no target-acquirer pair could do better by rematching, $\Sigma_t\left(z_t, \psi\left(z_t\right)\right) \geq 0$ for all z_t and $\Sigma_a\left(\psi^{-1}\left(z_a\right), z_a\right) \geq 0$ for all z_a, i.e., no firm prefers to remain standalone rather than enter its match. Taking the price of an acquisition as given, an acquiring firm chooses z_t to maximize its net of merger payment profits:

$$\max_{z_t} \; s\left(z_a, z_t\right) - P\left(z_t\right)$$

The first order condition gives

$$s_{z_t}\left(z_a, z_t\right) - \frac{dP\left(z_t\right)}{dz_t} = 0$$

and integrating on the equilibrium path where $z_a = \psi\left(z_t\right)$ gives the price[54]

$$P\left(z_t\right) = \int_{z_{min}}^{z_t} s_{z_t}\left(\tilde{z}_t, \psi\left(\tilde{z}_t\right)\right) d\tilde{z}_t$$

and the standard result of positive assortative matching, i.e,. $\psi\left(z\right) = z$, since targets are paid their marginal product and higher type acquirers have a higher marginal value for higher type targets and are willing to pay more for them. For the technology I consider, it is straightforward to show $P\left(z\right) = \frac{1}{2} z^{2\varphi}$.

Proof of sorting with repeat matching. Consider the case where $\varphi = 1$. From expression (11), a merger will go through when

$$\left(\frac{1}{2}z_t^2 - \frac{1}{2}\right) z_a^2 + \left(z_t - 1\right) z_a - \left(\frac{1}{2}z_t + 1\right) z_t \geq 0 \tag{25}$$

By definition, in the region of z_t where sorting can occur, i.e., where a firm is willing to match

[54] Because in this example the lowest types will not merge, the constant of integration will be zero.

with its own type, the lowest z's will match. But, expression (25) shows that the surplus is monotonically increasing in z_a (note that $z_t > 1$ for all z_t in the matching set), so there is also positive surplus from the lowest z_t matching with the highest z_a. So wherever there is matching, there is random matching, and sorting does not obtain. Now consider the case where $\varphi = \frac{1}{2}$. Here, the surplus condition collapses to the static one:

$$(z_a z_t)^{\frac{1}{2}} - z_a - z_t \geq 0$$

which never holds, so the matching set is empty. For values of $\varphi \in \left(\frac{1}{2}, 1\right)$, we must prove that the technology satisfies the conditions in Shimer and Smith (2000). Redefine the effective technology as $s(z_a, z_t) = (z_a z_t)^{\varphi} + \frac{1}{2}(z_a z_t)^{2\varphi^2}$ and the firms' standalone values as $z_a + \frac{1}{2} z_a^{2\varphi}$ and $z_t + \frac{1}{2} z_t^{2\varphi}$. It is straightforward to verify this technology is strictly supermodular. To prove the first partial derivative is log-supermodular, we need to show that for all $x_1 \leq x_2$ and $y_1 \leq y_2$, $s_x(x_1, y_1)\, s_x(x_2, y_2) \geq s_x(x_1, y_2)\, s_x(x_2, y_1)$. Taking the derivatives, the condition is

$$\left(\varphi x_1^{\varphi-1} y_1^{\varphi} + \varphi^2 x_1^{2\varphi^2-1} y_1^{2\varphi^2}\right)\left(\varphi x_2^{\varphi-1} y_2^{\varphi} + \varphi^2 x_2^{2\varphi^2-1} y_2^{2\varphi^2}\right) \geq$$
$$\left(\varphi x_1^{\varphi-1} y_2^{\varphi} + \varphi^2 x_1^{2\varphi^2-1} y_2^{2\varphi^2}\right)\left(\varphi x_2^{\varphi-1} y_1^{\varphi} + \varphi^2 x_2^{2\varphi^2-1} y_1^{2\varphi^2}\right)$$

and simplifying and rearranging gives

$$\left(x_1^{\varphi-1} x_2^{2\varphi^2-1} - x_1^{2\varphi^2-1} x_2^{\varphi-1}\right)\left(y_1^{\varphi} y_2^{2\varphi^2} - y_1^{2\varphi^2} y_2^{\varphi}\right) \geq 0$$

which always holds for $\varphi > \frac{1}{2}$. Similar steps give that the cross-partial derivative is log-supermodular, i.e., for all $x_1 \leq x_2$ and $y_1 \leq y_2$, $s_{xy}(x_1, y_1)\, s_{xy}(x_2, y_2) \geq s_{xy}(x_1, y_2)\, s_{xy}(x_2, y_1)$. Then the technology satisfies the conditions of Proposition 6 in Shimer and Smith (2000) and matching will be positive assortative (for the set of z willing to match with their own type).

$\square$

Proof of sorting without repeat matching. Without repeat matching, the surplus condition becomes:

$$2(z_a z_t)^{\varphi} - z_a - z_t - \frac{1}{2} z_a^{2\varphi} - \frac{1}{2} z_t^{2\varphi} \geq 0$$

For $\varphi = \frac{1}{2}$, this implies

$$2(z_a z_t)^{\frac{1}{2}} - \frac{3}{2}(z_a + z_t) \geq 0$$

which, by Jensen's Inequality, never holds (since $(z_a z_t)^{\frac{1}{2}} \leq \frac{1}{2}(z_a + z_t) \Rightarrow 2(z_a z_t)^{\frac{1}{2}} \leq z_a + z_t <$

$\frac{3}{2}(z_a + z_t))$. For $\varphi = 1$, the condition becomes

$$2z_a z_t - z_a - z_t - \frac{1}{2}z_a^2 - \frac{1}{2}z_t^2 \geq 0$$

which is a quadratic equation in z_a. We can verify that both roots are increasing in z_t, so sorting obtains.

$\square$

Proof of sorting when merged profits exceed standalone profits in all matches. I extend the production to be $z_a + z_t + (z_a z_t)^{\varphi}$. The surplus condition becomes

$$\begin{aligned} 0 \leq\ & z_a + z_t + (z_a z_t)^{\varphi} + \frac{1}{2}\left[2\left(z_a + z_t + (z_a z_t)^{\varphi}\right) + \left(z_a + z_t + (z_a z_t)^{\varphi}\right)^{2\varphi}\right] - z_a - z_t \\ -\ & \frac{1}{2}\left[2z_a + z_a^{2\varphi}\right] - \frac{1}{2}\left[2z_t + z_t^{2\varphi}\right] \end{aligned}$$

and after cancelling terms,

$$2(z_a z_t)^{\varphi} + \frac{1}{2}\left(z_a + z_t + (z_a z_t)^{\varphi}\right)^{2\varphi} - \frac{1}{2}z_a^{2\varphi} - \frac{1}{2}z_t^{2\varphi} \geq 0$$

which always holds. So all mergers in period 1 go through, i.e., there is random matching and sorting does not obtain.

Without repeat matching, which is the Shimer and Smith (2000)) setup, the surplus condition is

$$2(z_a z_t)^{\varphi} - \frac{1}{2}z_a^{2\varphi} - \frac{1}{2}z_t^{2\varphi} \geq 0$$

which is a quadratic equation in z_a^{φ} with roots given by $z_a = z_t\left(2 \pm \sqrt{3}\right)^{\frac{1}{\varphi}}$. So matching sets are increasing for any (finite) value of φ and sorting obtains. $\square$

C Estimation

C.1 Identification

Sorting in a one-shot example. To obtain a nonempty matching set, we need the upper bound to be greater than the lower bound for some z_t:

$$\frac{1}{1-\gamma}\log(\beta A) + \frac{\nu}{1-\gamma}\log z_t > -\frac{1}{\gamma}\log((1-\beta)A) + \frac{1-\nu}{\gamma}\log z_t$$

which can be arranged as

$$\log A + (1-\gamma)\log(1-\beta) + \gamma \log \beta > (1-\gamma-\nu)\log z_t$$

Assume that $A < \left(\beta^{\gamma}(1-\beta)^{1-\gamma}\right)^{-1}$.[55] Since $z_t > 1$ (it is straightforward to see that firms with $z_t < 1$ will not match and so fall outside the matching set) and the left-hand side is strictly negative, a necessary and sufficient condition for a nonempty matching set is $1-\gamma-\nu < 0$, or $\gamma + \nu > 1$. For the slopes to be positive, it clearly must be the case that $\gamma < 1$ and $\nu < 1$. Thus, necessary and sufficient conditions for sorting are (i) $\gamma + \nu > 1$ and (ii) $\gamma < 1, \nu < 1$.

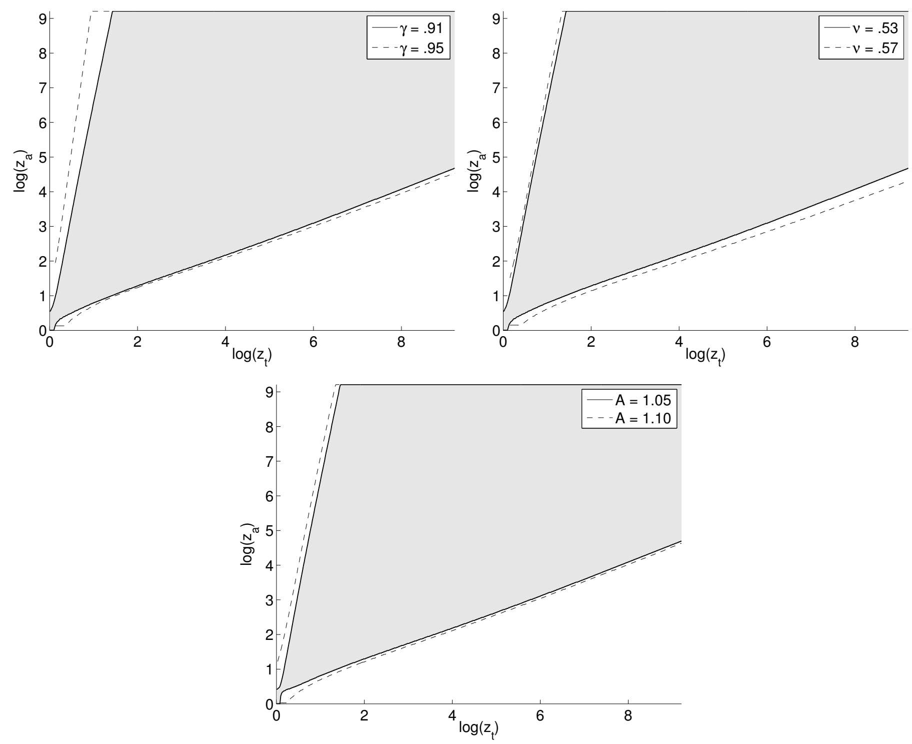
Numerical analog to Figure 4. Figure 9 provides a numerical analog in the full estimated model of the identification arguments in Figure 4. The top left panel shows the matching set at the estimated value of γ of 0.91 (shaded in grey) and for a higher value of 0.95 (the dashed lines), holding the other parameters fixed at their estimated values. The figure shows exactly the same patterns as Figure 4 - an increase in γ causes the matching set to rotate outward (the upper bound expands more than the lower). The top right panel shows the analogous result for an increase in ν - the matching set rotates outward, this time with a stronger affect on the lower bound. The bottom panel shows that an increase in A causes both bounds to shift and, further, that the effect is especially pronounced for low productivity firms.

Numeric identification. I use a series of figures to demonstrate the relationship between each parameter and the corresponding moment. Each figure is generated from the full estimated model by holding all the parameters fixed at their estimated values except the parameter of interest. I then vary that parameter around its estimated value and compute the new values of the associated moment.

Turning first to the parameters of the merger technology, Figure 10 plots values of γ, ν and A against the corresponding moments - the median acquirer type, the median target type and the share of small targets, respectively. The median acquirer type falls as γ increases - with less curvature, lower productivity firms have a larger number of profitable partners and so match more frequently. Following a similar logic, the median target type falls in ν. The bottom panel shows that as A increases, so does the share of small targets (those in the bottom decile) - with larger secular gains from merger (independent of firm types), less productive firms become more attractive targets.

Figure 11 illustrates the relationships between β, η and B and the corresponding moments - the mean merger premium, the coefficient of variation of target types and the rate of merger.

[55]The proof extends to larger values of A up to a relatively large upper bound (for example, if $\beta = \frac{1}{2}$, the bound is $A = 2$).



Notes:

Figure 9: Numeric Identification of Merger Gains

Each moment is sensitive to the corresponding parameter.

C.2 Invariance to Target Search Costs

Without loss of generality, write the cost of search for targets as BT, where B is the cost of search for acquirers and T is some positive number capturing the relative cost of search for targets. I prove that so long as $\theta_t < 1$, changes in target search costs, T, change market tightness, θ_t, and target search, $\mu(z)$, in exactly offsetting ways, keeping the total search costs and meeting rates of targets constant. Target search costs then have no other effects on the analysis, and so setting $T = 1$ is without loss of generality in this case. Following the proof, I show that this condition holds for a wide range of values for T.

Proof of invariance to T when $\theta_t < 1$. Consider the value function

$$rV(z) = \quad \pi(z) - C(\lambda(z)) - C(\mu(z)) + \lambda(z)\,\theta_a \mathbb{E}\left[\Sigma_a(z, z_t)\right] + \mu(z)\,\theta_t \mathbb{E}\left[\Sigma_t(z_a, z)\right]$$

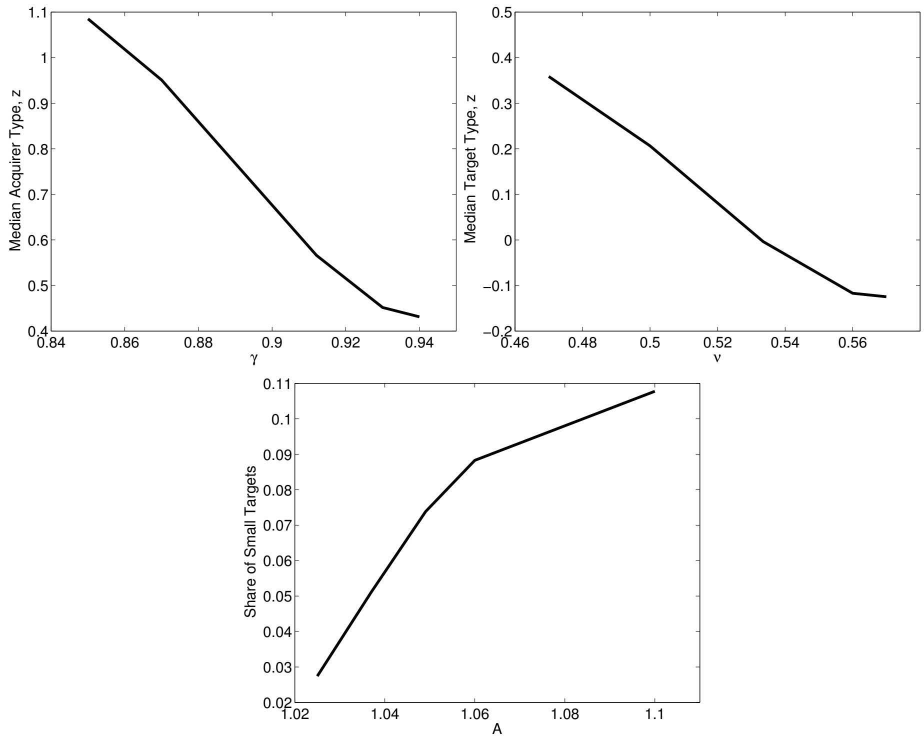


Figure 10: Merger Technology - Moments and Parameters

first-order conditions for search

$$\lambda(z) = \left(\frac{\theta_a \mathbb{E}\left[\Sigma_a(z, z_t)\right]}{B}\right)^{\frac{1}{\eta-1}}, \quad \mu(z) = \left(\frac{\theta_t \mathbb{E}\left[\Sigma_t(z_a, z)\right]}{BT}\right)^{\frac{1}{\eta-1}}$$

and steady state conditions

$$\begin{aligned} & M\int \lambda(z_a)\,\theta_a \Phi\left(\Sigma\left(z_a, s^{-1}(z, z_a)\right)\right)\Gamma(z_t)\,dG(z_a) + M_e dF(z) \\ = \; & \lambda(z)\,\theta_a M dG(z)\int \Phi\left(\Sigma(z, z_t)\right)\Gamma(z_t) \\ + \; & \mu(z)\,\theta_t M dG(z)\int \Phi\left(\Sigma(z_a, z)\right)\Lambda(z_a) + \delta M dG(z) \;\; \forall z \geq \hat{z} \end{aligned}$$

For $\theta_t < 1$, we have

$$\theta_t = \frac{\int \lambda(z)\,dG(z)}{\int \mu(z)\,dG(z)}$$

Without loss of generality, assume $T = 1$ and consider a change to $\hat{T}$. Conjecture that the value functions remain constant, which implies merger gains and matching sets remain constant.

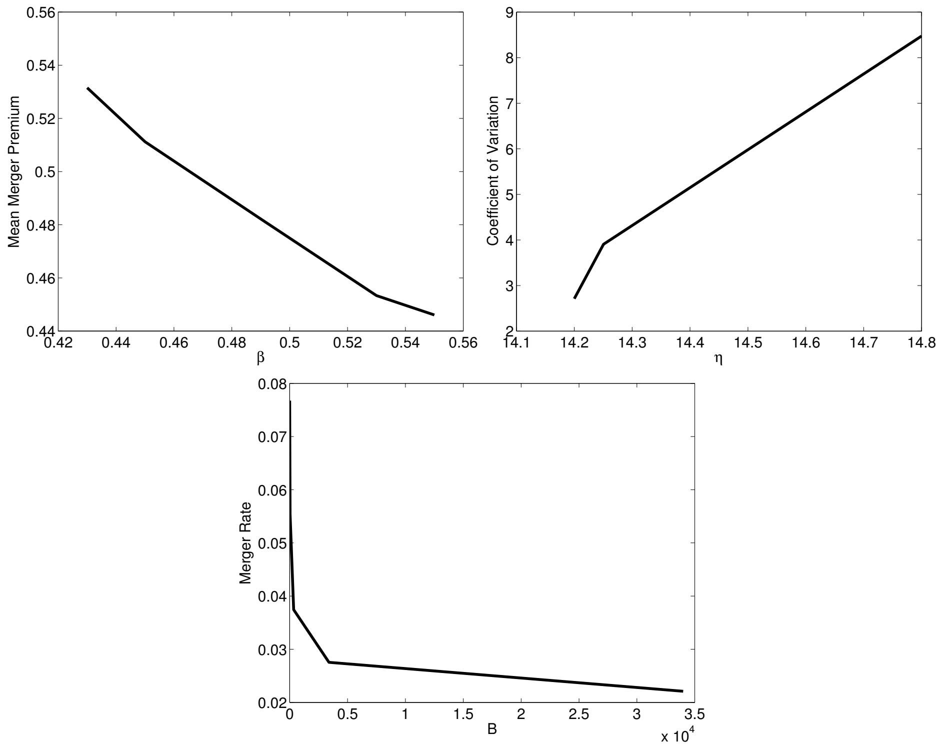


Figure 11: Merger Costs and Split - Moments and Parameters

Under this conjecture, $\lambda(z)$ is unchanged for all z. The new levels of $\hat{\mu}(z)$ and $\hat{\theta}_t$ must satisfy:

$$\hat{\mu}(z) = \left(\frac{\hat{\theta}_t \mathbb{E}\left[\Sigma_t(z_a, z)\right]}{B\hat{T}}\right)^{\frac{1}{\eta-1}}, \quad \hat{\theta}_t = \frac{\int \lambda(z)\, dG(z)}{\int \hat{\mu}(z)\, dG(z)}$$

It is straightforward to verify that these equations are satisfied for

$$\hat{\mu}(z) = \mu(z)\, \hat{T}^{-\frac{1}{\eta}}, \quad \hat{\theta}_t = \theta_t \hat{T}^{\frac{1}{\eta}} \tag{26}$$

Examining the terms involving target search in the value function, these are

$$\begin{aligned} -\frac{B\hat{T}}{\eta}\mu(z)^{\eta} + \mu(z)\,\theta_t \mathbb{E}\left[\Sigma(z_a, z)\right] &= -\frac{B\hat{T}}{\eta}\left(\mu(z)\,\hat{T}^{-\frac{1}{\eta}}\right)^{\eta} + \mu(z)\,\hat{T}^{-\frac{1}{\eta}}\theta_t \hat{T}^{\frac{1}{\eta}}\mathbb{E}\left[\Sigma(z_a, z)\right] \\ &= -\frac{B}{\eta}\mu(z)^{\eta} + \mu(z)\,\theta_t \mathbb{E}\left[\Sigma(z_a, z)\right] \end{aligned}$$

which are the same as in the original value function with $T = 1$. Turning to the steady state conditions, it is straightforward to verify that they are satisfied for the original distribution $dG(z)$: the first two lines are unchanged (since all $\mu(z)$ scale up or down proportionally, the distribution $\Gamma(z_t)$ is unaffected). The last line contains $\hat{\mu}(z)\hat{\theta}_t = \mu(z)\theta_t$, so is also unchanged, proving that the steady state distribution $dG(z)$ is unchanged. This verifies the conjecture that value functions remain constant. Because value functions, meeting rates and expenditures on search by targets are unaffected by the level of target search costs, they have no further effects on the analysis. $\square$

Intuitively, the proof shows that market tightness reacts endogenously to changes in T to keep target expenditures on search and meeting rates constant and so the analysis is invariant to the level of target search costs. Under the assumption of an equal level of costs on the two sides of the market, the estimated level of market tightness is $\theta_t = 0.86$. Expression (26) proves that the condition holds for any value of target search costs lower than acquirers' ($T < 1$), since lower levels of costs increase target search and reduce θ_t even further. We can also use expression (26) to derive an upper bound for which the condition holds - at an estimated θ_t of 0.86 and η of about 14, the condition is satisfied for target search costs up to about eight times that of acquirers'. Further, I have verified that with target search costs large enough such that the exact invariance condition does not hold, i.e., the short side of the market changes to $\theta_a < 1$ and $\theta_t = 1$, the parameter estimates change little - for example, assuming target search costs ten times that of acquirers yields estimates quite close to the baseline.

C.3 Alternative Approaches

Match-specific shocks. I consider an extension of the technology in which the gains from a match depend on a match-specific component ε, so that the technology is $\hat{z}_m = \hat{A} + \gamma\hat{z}_a + \nu\hat{z}_t + \varepsilon_{a,t}$ where hats denote natural logs. If this component is unobserved, then clearly it cannot affect firms' matching choices. In this case, my strategy goes through, as does the statistical approach. Now assume that the firm observes a signal of this shock, κ, but which is unobserved by the econometrician. Taking a log-linear approximation to the value function, $V(z_i) \approx \phi\hat{z}_i$, and denoting the decision to merge by m, we have

$$m = 1 \text{ iff } \mathbb{E}\left[\hat{z}_m - \hat{z}_a - \hat{z}_t | \kappa\right] \geq 0 \quad \Rightarrow \quad \hat{A} + (\gamma - 1)\hat{z}_a + (\nu - 1)\hat{z}_t + \frac{\sigma_{\varepsilon,\kappa}}{\sigma_\kappa^2}\kappa \geq 0$$

from which we can show

$$\mathbb{E}\left[\hat{z}_m | m = 1\right] = \hat{A} + \gamma\hat{z}_a + \nu\hat{z}_t + \frac{\sigma_{\varepsilon,\kappa}}{\sigma_\kappa}\lambda\left(\frac{\hat{A} + (\gamma - 1)\hat{z}_a + (\nu - 1)\hat{z}_t}{\sigma_\kappa}\right) \tag{27}$$

where λ denotes the inverse Mills ratio, which is a model of sample selection. The conditional expectation of $\hat{z}_m$ is not equal to the true technology and estimating the associated regression is subject to an omitted variable bias, which would tend to bias the coefficients downward. Indeed, estimating (27) on my sample and ignoring the selection term yields coefficients of around 0.8 and 0.1, both highly significant, but well below the baseline estimates, consistent with the presence of downward bias. In principle, one could include the correction term, but the situation here is complicated, since the observables determining selection (z_a and z_t) are exactly those that determine the outcome, z_m, which is likely to introduce multicollinearity into the regression. Li and Prabhala (2007) recommend using additional instruments that affect the selection decision but are independent of outcomes following selection, however, it is not clear what suitable instruments might be.[56] Thus, identifying the technological parameters off the statistical relationship between the pre- and post-merger firms' performance is potentially quite challenging.

Next, I explore whether these same concerns are likely to substantially bias my estimates. I simulate a dataset of matched acquirers and targets, z_a and z_t, where $\log z \sim \mathcal{N}(0, \sigma_z^2)$ and a set of match-specific shocks ε, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, which I assume firms observe perfectly. I set $\sigma_z^2 = 1$ and $\sigma_\varepsilon^2 = 0.1$. The merger technology is $z_m = A z_a^\gamma z_t^\nu e^\varepsilon$ where $A = 1.05, \gamma = 0.93, \nu = 0.53$, the estimated values from the text. Firms merge only if this exceeds $z_a + z_t$. I then select the transactions that satisfy this condition and regress $\log z_m$ on $\log z_a$ and $\log z_t$. The first row of Table 15 shows that the presence of selection leads to a substantial downward bias in the parameter estimates - the estimated values are 0.81 and 0.46, compared to the true values of 0.91 and 0.53. Next, I use this same simulated dataset to estimate the parameters using the same approach as in the main text, i.e., targeting the median z_a and z_t, but importantly, using the misspecified technology $z_m = A z_a^\gamma z_t^\nu$. The second row of Table 15 shows that doing so gives estimates that are much closer to the true values, 0.94 and 0.53 (and exact in the case of ν).

Table 15: Match-Specific Shocks - Potential Bias

	γ		ν	
Estimation Approach	True	Estimated	True	Estimated
Pre/Post-Merger Regression	0.91	0.81	0.53	0.46
Pre-Merger Sorting	0.91	0.94	0.53	0.53

Regression-based evidence of performance improvements. I follow the approach in, e.g., Maksimovic and Phillips (2001) and estimate regressions of the following forms

$$x_{it} = \beta_0 + \beta_1 x_{it-3} + \Phi\left(acq_{it-2} = 1\right) + D_{jt} + \zeta_{it}$$

[56] An additional issue is that the equation above only holds when errors are normal. If not, even the selection model will be biased.

where x_{it} denotes a measure of firm i performance in year t, $\Phi(\cdot)$ denotes an indicator variable equal to one if the firm made an acquisition two years prior and D_{jt} denotes a full set of industry-year fixed effects. This specification asks whether the performance of acquiring firms is affected by acquisition activity, controlling for its pre-merger performance and relative to the other firms in its industry and year. First, I examine the acquiring firm alone and construct two scale-independent measures of performance - labor productivity, calculated as log revenues less log employment and total factor productivity, calculated as $\log Y - \frac{1}{3}\log K - \frac{2}{3}\log L$. Second, I directly examine the effects on size/profitability. To do so, I construct a synthetic level of pre-merger size/profitability of the combined firm by adding together the revenues/profits of the two pre-merger firms in the year before the merger. I then estimate the regression using the post-merger firm's sales/profits as the left-hand side variable and the the combined pre-merger sales/profits as the right-hand side variable.

I report the results in Table 16. Columns (1) and (2) show that both acquirer total factor productivity and labor productivity increase relative to their pre-merger values upon an acquisition. The coefficients on the acquisition indicator are both positive and statistically significant at standard levels. Similarly, columns (3) and (4) show that the sales and profitability of the combined entity increase significantly relative to the sum of the standalone firms.[57] Both sets of results point to the presence of significant performance gains surrounding a merger event.

C.4 Random Search

I consider the set of transactions for which I have data on profitability for both the acquirer and target from SDC. This is about 5,000 transactions. I allocate acquirers into deciles ranked by profitability, i.e., decile 1 is the least profitable 10% of acquirers and decile 10 the most profitable. Next, within each acquirer decile, I allocate their targets into deciles, so that, for example, group $(1,1)$ contains the 10% least profitable targets of the 10% least profitable acquirers, group $(1,2)$ the 10% second-least profitable targets of the least profitable acquirers, up to group $(10,10)$, which contains the most profitable targets of the most profitable acquirers. There are 100 such groups, with about 50 transactions per group. For each group, I calculate the bounds of the matching set, i.e., the minimum and maximum target types. Using the intervals of targets defined by these bounds, I perform the KS-test relative to the same interval of targets for the next highest acquirer decile. I conduct the test for all target intervals defined by target deciles 6-10. For example, beginning with acquirer decile 1, I find the upper and lower

[57]The coefficients suggest a larger effect on these measures than on the measures of productivity. This result echoes that in Braguinsky et al. (2015), who use an extremely detailed dataset to show that there are substantial gains upon merger, but that gains may manifest themselves more strongly in broader measures of firm performance.

Table 16: Pre- and Post-Merger Performance

	TFP (1)	LP (2)	Sales (3)	Profitability (4)
acq_{t-2}	0.021***	0.041***	0.181***	0.210***
	(0.005)	(0.005)	(0.015)	(0.019)
x_{t-3}	0.620***	0.661***	0.918***	0.910***
	(0.002)	(0.002)	(0.001)	(0.002)
Constant	1.548***	1.839***	0.630***	0.515***
	(0.008)	(0.010)	(0.006)	(0.006)
Observations	119097	122564	110344	82809
R^2	0.69	0.76	0.91	0.88
Industry-Year Effects	Yes	Yes	Yes	Yes

Notes: Table reports regressions of post- on pre-merger firm performance. x denotes a measure of firm-level performance in year t. Columns (1) and (2) show estimates using total factor productivity and labor productivity of the acquiring firm. Columns (3) and (4) show estimates using the sales and profitability of the merged firm and the sum of the two pre-merger firms. Standard errors are in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

bounds for targets that are within decile 6 (group $(1, 6)$). These are about \$1.8 million and \$2.3 million in profits, respectively (in 2005 dollars). I then test the equality of the distribution of targets that fall in this interval for acquirers in decile 1 compared to decile 2. I repeat this exercise for all such groups, e.g., for the interval of targets defined by acquirer decile 1, target decile 6, up to the last group, the interval of targets defined by acquirer decile 9, target decile 10 (where I test the equality of the distribution of these targets purchased by acquirers in decile 9 to acquirers in decile 10). This gives a total of 45 groups. For 43 of them, the KS-test fails to reject the null hypothesis that the distributions are equal at the 95% confidence level. The results are similar if I use sales as a measure of target type (for which I have about 12,200 observations). The results of the 45 tests are reported in Online Appendix E.

To give a visual sense of what the tests are capturing, Figure 12 shows two examples comparing the densities of overlapping targets across two groups of acquirers. The left-hand plot shows the interval of targets defined by the bounds of the 9th target decile of the 5th aquirer decile and compares the density over this group of targets for acquirers in deciles 5 and 6. This is an example where the null of equality cannot be rejected. Visually, the densities are very close. The right-hand panel examines the interval of targets defined by the bounds of the 7th target decile of the 2nd acquirer decile and compares the density over this group of targets for acquirers in deciles 2 and 3. This is one of the two pairings for which the null of equivalent distributions can be rejected at the 95% confidence level. The first density has much of its mass in the left tail and is decreasing in target type, whereas the second density is more evenly distributed across the range of targets.

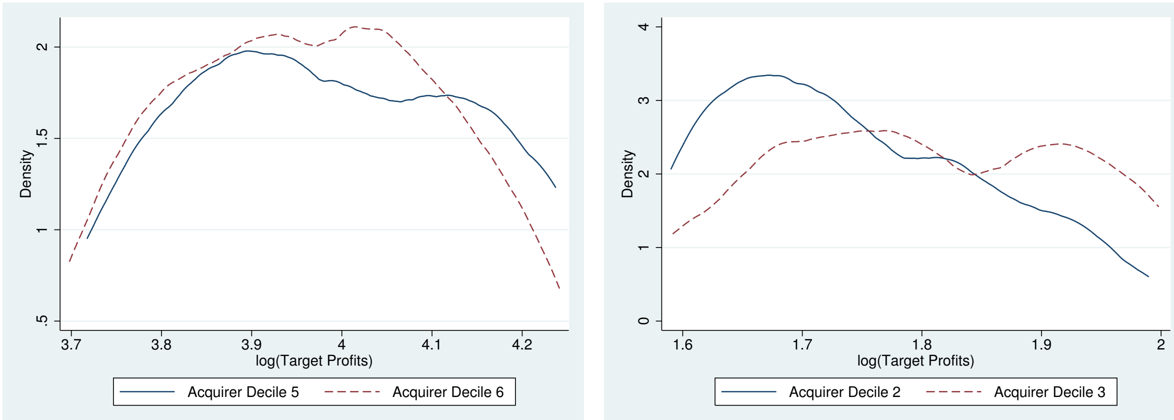


Figure 12: Example Density Plots of Common Targets Across Acquirer Types

D Efficiency

Static allocation. We can split the planner's problem into two components, the static labor allocation problem, taking as given the distribution of firms, and the dynamic search and matching problem which in part determines that distribution. The static problem is to maximize total production subject to the feasibility constraint on labor:

$$\begin{aligned} \max_{l(z)} \quad & M \int z^{1-\alpha} l(z)^{\alpha} \, dG(z) \\ & s.t. \\ & M \int l(z) \, dG(z) = L \end{aligned}$$

Solving this gives $l(z) = \frac{z}{M\overline{Z}} L$ and integrating gives the same aggregate production function as in (15). The contribution of marginally increasing the density of type z $(MdG(z))$ to aggregate output is given by

$$\frac{d}{dMdG(z)} \left(\int MzdG(z) \right)^{1-\alpha} L^{\alpha} = (1-\alpha) \left(M\overline{Z} \right)^{-\alpha} L^{\alpha} z$$

Imposing labor market clearing on expression (12), it is straightforward to verify that $l(z) = \frac{z}{M\overline{Z}} L$ and $\Pi z = (1-\alpha) \left(M\overline{Z} \right)^{-\alpha} L^{\alpha} z$. So the allocation of labor in the decentralized economy is the same as the planned one, and the level of profits of type z are that type's contribution to aggregate output.

Dynamic problem. With these results in hand, we can write the planner's dynamic problem as

$$\begin{aligned}
\max_{\substack{\lambda_s(z),\mu_s(z),\\ \Phi_s(\Sigma(z_a,z_t))\ \forall\ z_a,z_t}} & \int_0^\infty e^{-rs}\left\{M_s\int\left[\Pi_s z-C\left(\lambda_s\left(z\right)\right)-C_s\left(\mu\left(z\right)\right)-c_f\right]dG_s\left(z\right)-M_{e,s}c_e\right\}ds \\
& s.t. \\
M_s d\dot{G}_s\left(z\right) &= M_{e,s}dF\left(z\right)-\delta M_s dG_s\left(z\right) \\
&+ M_s\theta_{a,s}\int\lambda_s\left(z_a\right)\Phi_s\left(\Sigma\left(z_a,\tilde{z}_t\right)\right)\frac{\mu_s\left(\tilde{z}_t\right)dG_s\left(\tilde{z}_t\right)}{\int\mu_s\left(z\right)dG_s\left(z\right)}dG_s\left(z_a\right) \\
&- \lambda_s\left(z\right)\theta_{a,s}M_s dG_s\left(z\right)\int\Phi\left(\Sigma\left(z,z_t\right)\right)\frac{\mu_s\left(z_t\right)dG_s\left(z_t\right)}{\int\mu_s\left(z\right)dG_s\left(z\right)} \\
&- \mu_s\left(z\right)\theta_{t,s}M_s dG_s\left(z\right)\int\Phi\left(\Sigma\left(z_a,z\right)\right)\frac{\lambda_s\left(z_a\right)dG\left(z_a\right)}{\int\lambda_s\left(z\right)dG\left(z\right)} \\
& \text{where } \tilde{z}_t=\left\{z_t: s\left(z_a,z_t\right)=z\right\}
\end{aligned}$$

where subscript s denotes the dependence on time. The planner chooses the time paths of search intensities and acceptance regions to maximize the discounted present value of consumption, which is equal to output less costs of search less fixed and entry costs, subject to a sequence of dynamic constraints regulating the change in the densities of firms. The constraint shows that the change in the density of a type z firm, $MdG\left(z\right)$, is equal to new entry of that type less exit (the first row) plus the flow in of new of firms through merger (the second row) less the flow out through merger (either as a target or acquirer).

We can write the current-valued Hamiltonian as

$$\begin{aligned}
H &= M\int\left[\Pi z-C\left(\lambda\left(z\right)\right)-C\left(\mu\left(z\right)\right)-c_f\right]dG\left(z\right)-M_e c_e \\
&+ \int W\left(z\right)M\theta_a\int\lambda\left(z_a\right)\Phi\left(\Sigma\left(z_a,\tilde{z}_t\right)\right)\frac{\mu\left(\tilde{z}_t\right)dG\left(\tilde{z}_t\right)}{\int\mu\left(z\right)dG\left(z\right)}dG\left(z_a\right)dz \\
&- \int W\left(z\right)\lambda\left(z\right)\theta_a M dG\left(z\right)\int\Phi\left(\Sigma\left(z,z_t\right)\right)\frac{\mu\left(z_t\right)dG\left(z_t\right)}{\int\mu\left(z\right)dG\left(z\right)}dz \\
&- \int W\left(z\right)\mu\left(z\right)\theta_t M dG\left(z\right)\int\Phi\left(\Sigma\left(z_a,z\right)\right)\frac{\lambda\left(z_a\right)dG\left(z_a\right)}{\int\lambda\left(z\right)dG\left(z\right)}dz \\
&+ \int W\left(z\right)\left(M_e dF\left(z\right)-\delta M dG\left(z\right)\right)dz
\end{aligned}$$

where I have suppressed time subscripts and $W\left(z\right)$ denotes the multiplier on the constraint for type z. Notice that

$$M\theta_a\int\lambda\left(z_a\right)\Phi\left(\Sigma\left(z_a,z_t\right)\right)\frac{\mu\left(z_t\right)dG\left(z_t\right)}{\int\mu\left(z\right)dG\left(z\right)}dG\left(z_a\right)=$$

$$M\theta_t \int \mu(z_t)\,\Phi(\Sigma(z_a, z_t))\,\frac{\lambda(z_a)\,dG(z_a)}{\int \lambda(z)\,dG(z)}dG(z_t)$$

so that

$$\begin{aligned}
&\int W(z)\,M\theta_a \int \lambda(z_a)\,\Phi(\Sigma(z_a, \tilde{z}_t))\,\frac{\mu(\tilde{z}_t)\,dG(\tilde{z}_t)}{\int \mu(z)\,dG(z)}dG(z_a)\,dz = \\
&\frac{1}{2}\int W(z)\,M\theta_a \int \lambda(z_a)\,\Phi(\Sigma(z_a, \tilde{z}_t))\,\frac{\mu(\tilde{z}_t)\,dG(\tilde{z}_t)}{\int \mu(z)\,dG(z)}dG(z_a)\,dz + \\
&\frac{1}{2}\int W(z)\,M\theta_t \int \mu(z_t)\,\Phi(\Sigma(\tilde{z}_a, z_t))\,\frac{\lambda(\tilde{z}_a)\,dG(\tilde{z}_a)}{\int \lambda(z)\,dG(z)}dG(z_t)\,dz
\end{aligned}$$

where $\tilde{z}_a$ is defined analogously to $\tilde{z}_t$, i.e., $\tilde{z}_a = \{z_a : s(z_a, z_t) = z\}$.

Making this substitution on the second line of the Hamiltonian, the steady state costate equation gives (see Online Appendix F)

$$\begin{aligned}
(r+\delta)\,W(z) &= \pi(z) - C(\lambda(z)) - C(\mu(z)) && (28)\\
&+ \lambda(z)\,\theta_a\mathbb{E}[\Sigma(z, z_t)] + \mu(z)\,\theta_t\mathbb{E}[\Sigma(z_a, z)] \\
&- \mu(z)\,\theta_t\mathbb{E}[\Sigma(z_a, z_t)] \quad \text{if } \theta_t < 1 \text{ and } \theta_a = 1 \\
&- \lambda(z)\,\theta_a\mathbb{E}[\Sigma(z_a, z_t)] \quad \text{if } \theta_t = 1 \text{ and } \theta_a < 1 \\
&- \lambda(z)\,\theta_a\mathbb{E}\left[\frac{W(z_m)}{2} - W(z_a)\right] - \mu(z)\,\theta_t\mathbb{E}\left[\frac{W(z_m)}{2} - W(z_t)\right] \\
&\quad \text{if } \theta_t = 1 \text{ and } \theta_a = 1
\end{aligned}$$

which, in the case of $\theta_t < 1$, is expression (21) in the text.

The first order conditions for $\lambda(z)$ and $\mu(z)$ give (see Online Appendix F)

$$\begin{aligned}
c'(\lambda(z)) &= \theta_a\mathbb{E}[\Sigma(z, z_t)] && (29)\\
&- \theta_a\mathbb{E}[\Sigma(z_a, z_t)] \quad \text{if } \theta_t = 1 \text{ and } \theta_a < 1 \\
&- \theta_a\mathbb{E}\left[\frac{W(z_m)}{2} - W(z_a)\right] \quad \text{if } \theta_t = 1 \text{ and } \theta_a = 1 \\
c'(\mu(z)) &= \theta_t\mathbb{E}[\Sigma(z_a, z)] && (30)\\
&- \theta_t\mathbb{E}[\Sigma(z_a, z_t)] \quad \text{if } \theta_t < 1 \text{ and } \theta_a = 1 \\
&- \theta_t\mathbb{E}\left[\frac{W(z_m)}{2} - W(z_t)\right] \quad \text{if } \theta_t = 1 \text{ and } \theta_a = 1
\end{aligned}$$

which, in the case of $\theta_t < 1$, is equation (22). Finally, the first order condition on M_e gives expression (23) and the entry threshold will be set at $\hat{z}$, where $\hat{z}$ satisfies $W(\hat{z}) = 0$.

Under a system of linear taxes, the equilibrium value function satisfies

$$\begin{aligned} rV(z) = \max_{\lambda(z),\mu(z)} & \quad \pi(z) - \tau^{\lambda}\lambda(z) - \tau^{\mu}\mu(z) - C(\lambda(z)) - C(\mu(z)) \\ & + \quad \lambda(z)\,\theta_a\mathbb{E}\left[\Sigma_a(z, z_t)\right] + \mu(z)\,\theta_t\mathbb{E}\left[\Sigma_t(z_a, z)\right] \end{aligned}$$

It is straightforward to show that the following tax system decentralizes the social optimum:

$$\begin{aligned} \tau^{\lambda} &= -(1-\beta)\,\theta_a\mathbb{E}\left[\Sigma(z, z_t)\right] \\ &+ \quad \theta_a\mathbb{E}\left[\Sigma(z_a, z_t)\right]] \quad \text{if } \theta_t = 1 \text{ and } \theta_a < 1 \\ &+ \quad \theta_a\mathbb{E}\left[\frac{W(z_m)}{2} - W(z_a)\right] \quad \text{if } \theta_t = 1 \text{ and } \theta_a = 1 \\ \tau^{\mu} &= -\beta\theta_t\mathbb{E}\left[\Sigma(z_a, z)\right] \\ &+ \quad \theta_t\mathbb{E}\left[\Sigma(z_a, z_t)\right]] \quad \text{if } \theta_t < 1 \text{ and } \theta_a = 1 \\ &+ \quad \theta_t\mathbb{E}\left[\frac{W(z_m)}{2} - W(z_t)\right] \quad \text{if } \theta_t = 1 \text{ and } \theta_a = 1 \end{aligned}$$