Entry by Merger:

Estimates from a Two-Sided Matching Model with Externalities*

Kosuke Uetake† Northwestern University

Yasutora Watanabe‡ Northwestern University

August 2012

Abstract

As firms often acquire incumbents to enter a new market, presence of desirable acquisition targets affect both merger and entry decisions simultaneously. We study these decisions jointly by considering a two-sided matching model with externalities to account for the "with whom" decision of merger and to incorporate negative externalities of post-entry competition. By estimating this model using data on commercial banks, we investigate the effect of the entry deregulation by the Riegle-Neal Act. After proposing a deferred acceptance algorithm applicable to the environment with externalities, we exploit the lattice structure of stable allocations to construct moment inequalities that partially identify the banks' payoff function including potential (dis)synergies. We find greater synergies between larger and healthier potential entrants and smaller and less-healthy incumbent banks. Compared with de novo entry, entry barriers are much lower for entry by merger. By prohibiting de novo entry, our counterfactual quantifies the effect of the deregulation.

Keywords: Entry, merger, two-sided matching, partial identification, Riegle-Neal Act

*We thank Hiroyuki Adachi, David Besanko, Ivan Canay, Jeremy Fox, John Hatfield, Marc Henry, Hide Ichimura, Michihiro Kandori, Mike Mazzeo, Aviv Nevo, Rob Porter, Mark Satterthwaite, Matt Shum, Junichi Suzuki, Jeroen Swinkels, Rakesh Vohra, Yosuke Yasuda, and especially Fuhito Kojima, Isa Hafalir, and Elie Tamer, for their valuable comments and suggestions. We also thank Xiaolan Zhou for kindly sharing her data on bank mergers.

†Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208. Email: uetake@northwestern.edu

‡Department of Management and Strategy, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208. Email: y-watanabe@kellogg.northwestern.edu

1 Introduction

Firms often use mergers and acquisitions to enter new markets.[1] In such cases, presence of desirable acquisition targets affects not only merger decisions but also entry decisions at the same time. For example, a firm may choose not to enter a market if it cannot find a good target incumbent for acquisition. In some markets, entry barriers for *de novo* entry can be so high that acquiring an incumbent may be the only profitable way to enter. Thus, entry and merger decisions are joint decisions, and should not be separately studied in those markets. Moreover, as entry by merger impacts the competitive structure of markets, studying entry by merger may have important policy implications.[2]

In this paper, we study entry and merger decisions jointly in the U.S. commercial banking industry where entry by merger is prevalent. In particular, we investigate the effect of the Riegle-Neal Act (the Act), which deregulated the *intra*-state de novo entry for 13 states that had not deregulated it at the time of the Act.[3] In these 13 states, the legal barrier for de novo entry was eliminated by the Act in 1997. Using data on bank behavior in the regional markets of these 13 states for the period right after the Act took effect, we study how the Act affected entry and merger decisions.

To study the banks' behavior, we consider a model in which the "with whom" decision of the merger and the post-entry (and post-merger) competition are addressed simultaneously. We do so by combining a standard entry model (Bresnahan and Reiss, 1991, and Berry, 1992) with a two-sided matching model with contracts (Hatfield and Milgrom, 2005). For the "with whom" decision, two features are particularly important: i) the payoff from merger depends on potential (dis)synergy, which significantly differs across pairs of firms (we specify a *synergy function* to quantify potential (dis)synergies), and ii) the merger decision is not a unilateral one because the target firm must agree to the merger contract. Reflecting these features, we adopt a matching model. Specifically, we consider a one-to-one two-sided matching model in which firms are partitioned into two sides, incumbents ($\mathcal{I}$) and

[1] A significant fraction of market entry is reported to be by merger. Yip (1982) reported that more than one-third of entries in 31 product markets in the United States over the period 1972-1979 were by acquisition. Among 558 market entries into the United States by Japanese companies during 1981-1989, Hennart and Park (1993) found that entry by merger accounted for more than 36%.

[2] Enty by merger has been an important antitrust issues. The Federal Trade Commission (FTC) explicitly considers "potential new competitors" in its Merger Guidelines. A classic case is the FTC's decline of a merger attempt by Procter & Gamble (P&G) and the Crolox Corporation in 1967. FTC argued that P&G was the most likely potential entrant in the household bleach industry, and that P&G's acquisition of Clorox would eliminate P&G as a potential competitor, which would substantially reduce the competitiveness of the industry.

[3] The Act deregulated both *inter*- and *intra*-state banking regulations. Historically, there had been tight state-level regulations on inter- and intra-state de novo entries. After a period of gradual state-level deregulation, the Riegle-Neal Act removed the regulation on both inter- and intra-state de novo entry at the federal level in 1997. The other 37 states have already deregulated the intra-state de novo entry by the time of the Act.

potential entrants ($\mathcal{E}$), because the vast majority of the mergers in our data are mergers between one incumbent and one potential entrant.[4]

Regarding how we incorporate the effect of post-entry competition into the matching model, we follow a standard entry model (Bresnahan and Reiss, 1991, and Berry, 1992) in which the effect of competition on profit is modeled as a decreasing function of the number of operating firms. Considering this effect of competition on profit (negative externalities), a firm chooses the best option out of the three types of options {*Enter with merger, Enter without merger, Do not enter*}.[5] Because the profit of a firm depends not only on the firm's matching (merger partner) but also on the merger and entry decisions of other firms (e.g., whether another incumbent is acquired by a potential entrant), the model we consider becomes a two-sided matching model *with externalities*.

Considering externalities in a matching model poses a challenge (see, e.g., Sasaki and Toda, 1996, and Hafalir, 2008). This is because, with externalities, payoffs depend not only on matching but also on the entire assignment of who match with whom. The solution concept in such a case thus has to take into consideration, for each deviation, what the entire assignment would be in addition to with whom the deviating players would match. To incorporate what the assignment would be after each deviation, Sasaki and Toda (1996) and Hafalir (2008) propose an *estimation function* that maps each deviation to an expectation about the possible assignments following the deviation. Despite taking this approach, the model with general form of externalities is complex and the existence of a stable matching requires strong conditions. In our case, however, complexities are reduced by the fact that the externalities take a particular form — it depends only on the aggregate number of operating firms negatively (i.e., negative network externalities). Hence, we can modify the estimation function approach so that the estimation is only about the aggregate number of operating firms.

As we incorporate externalities, the definition of stability as a solution concept has to be modified accordingly. In addition to the regular definition of stability (i.e., individual rationality and no-blocking-pair condition), we require what we call *consistency of estimation*: The estimated number of operating firms should be consistent with the actual number of operating firms. Because we can show that the number of actual operating firms is decreasing function of the estimated number of operating firms (i.e., as the estimation on the number of operating firms increases, less firms have incentive to enter.), we can prove

[4] This fact reflects the types of market we observe: our data is from small regional markets where average number of incumbents are less than 5 (see Section 3 for the detail), thus mergers between incumbents or mergers involving more than two incumbents tend to be infeasible due to antitrust concerns.

[5] As we consider a matching model, the first option requires the consent of the matching partner for the matching to be stable, while the last two options can be chosen unilaterally (See Section 2.2.2 for more detail). Our model differs from a standard two-sided matching model in that there are two outside options, i.e., firms can be unmatched in two ways by entering without merger or by not entering the market.

the existence of a solution. We do so by proposing a generalized Gale-Shapley algorithm with externalities. The algorithm is a nested fixed point algorithm that has Hatfield and Milgrom's (2005) generalized Gale-Shapley algorithm as an inner loop conditional on the estimated number of operating firms in the market. The outer loop searches for the fixed point of the estimated number of operating firms so that it satisfies the consistency of estimation.[6]

Another challenge is an econometric one: the model has multiple equilibria and the parameter cannot be point-identified without imposing some equilibrium selection rules. Instead of imposing equilibrium selection rules, we take the partial-identification approach exploiting the lattice property of the equilibrium. Though the econometrician cannot tell which equilibrium is played in the observed data, the equilibrium characterization provides upper and lower bounds for the equilibrium payoffs for each firm. To be more precise, all incumbents have the highest equilibrium payoff in incumbent($\mathcal{I}$)-optimal equilibrium, and the lowest equilibrium payoff in potential entrant($\mathcal{E}$)-optimal equilibrium. All other equilibrium payoffs are bounded by these two. Hence, the payoff corresponding to the observed outcome is bounded above and below by these extremum equilibrium payoffs, from which we construct moment inequalities for incumbents. Similarly, all potential entrants obtain the highest payoff in $\mathcal{E}$-optimal equilibrium and so on. Thus, we can construct moment inequalities using these equilibrium characterizations without the knowledge of the equilibrium selection rule.

The identified set can be reduced further by considering other equilibrium properties on top of the moment inequalities in payoffs. In all equilibria, the model predicts a unique number of operating firms and a unique number of mergers (the result is analogous to the *lone wolf theorem* and the *rural hospitals theorem*. See, e.g., Roth, 1986, and Hatfield and Milgrom, 2005). Using these properties, we construct moment equalities on the number of operating firms in addition to the moment inequalities on payoffs.

Identification of the model, including the identification of the synergy function, is obtained by two types of exclusion restrictions. The first type of exclusion restriction is a standard one used in the entry model literature following Berry (1992) and Ciliberto and Tamer (2009): we consider a variable that affects a firm's own payoff but not the other firms' profits. The second type of exclusion restriction takes advantage of match-specific characteristics: we consider a variable that influences only the (dis)synergy between the pair of firms, but do not influence the profits of the firms of the pair if they enter without merger and the (dis)synergy of any other combination of firms. Using the second exclusion

[6] Because the number of operating firms in the market monotonically decreases the profit of all firms (i.e., negative externalities of competition), the algorithm always converges. Furthermore, we can show that the convergence is global because of quasi-concavity of the objective function we minimize in the outer-loop of the algorithm.

restriction, we can make any merger to be less attractive than not entering. Then, the model become equivalent to the regular entry model because entering without merger and not entering are the only viable choices in such cases. Thus, we can use the first exclusion restriction to identify payoffs for entering without merger in the same way as the regular entry model is identified. Now, given the payoff of entering without merger is identified, the synergy function can be identified from the variations in merger probabilities and in firm characteristics.

Based on the moment equalities and inequalities discussed above, we estimate the model using Andrews and Soares's (2010) generalized moment selection. In computing the sample analogue of the moments, we run the generalized Gale-Shapley algorithm with externalities in order to obtain both $\mathcal{I}$-optimal and $\mathcal{E}$-optimal equilibria for each simulation draw for each market.

We find a significant difference in entry barriers for incumbents, potential entrants, and entry by merger. The entry barrier for potential entrants is much higher than that of the incumbents as well as that of entry by merger, causing a significant fraction of entry by potential entrants to take the form of merger and acquisition. Concerning the (dis)synergies from mergers between different types of firms, bank characteristics affect (dis)synergy differently across the sides of incumbents and potential entrants. We find that synergy between potential entrants with a larger asset size and higher equity ratio and incumbents with a lower equity ratio tends to be much higher. This may reflect the pattern in the data that large and healthy banks enter new markets by buying incumbents with less-healthy balance sheets.

Finally, we conduct a counterfactual policy experiment to assess the effect of the Act. We find the number of banks operating in a market would have decreased if de novo entry by potential entrants had not been deregulated by the Act. Prohibition of de novo entry provides stronger incentive to enter by merger for potential entrants, and accordingly we find that the number of entry by merger increases if de novo entry were prohibited.

In the rest of the paper, we present a two-sided matching model in Section 2 after discussing the related literature in Section 1.1. All proofs for Section 2 are in Appendix A. We document the data in Section 3, and then provide the econometric specification, identification, and estimation procedure in Section 4. Section 5 reports the results of the estimation and the counterfactual experiment. Finally, we conclude in Section 6.

1.1 Related Literature

Our paper adds to several strands of literature. The first strand of the related literature is the literature on estimating matching models. We add to this literature by proposing

a new approach to estimate a matching model with *non-transferrable* utility.[7] Only a few papers estimate matching models with non-transferrable utility. Gordon and Knight (2009) consider merger of school districts, and their specification on match quality is similar to our synergy function, though they consider a one-sided matching model (roommate problem). Also their paper is close to ours as they run an algorithm to find a stable matching in their estimation. Sorensen (2007) studies the matching between venture capitalists and entrepreneurs, and estimates the model under the assumption that players have aligned preferences. Boyd et al. (forthcoming) similarly estimate a two-sided matching model between teachers and schools by running the Gale-Shapley algorithm with the assumption that the school-optimal stable matching is realized. Similarly with a known equilibrium selection mechanism, Uetake and Watanabe (2012) propose another estimation strategy for a two-sided matching model with non-transferable utility using Adachi's (2000) prematching mapping. In environments with only aggregate-level data available, Echenique, Lee, Shum, and Yenmez (forthcoming) study testable implications of stable matchings. In a similar environment, Hsieh (2011) proposes a modified deferred acceptance algorithm to study identification and estimation. These two papers differs from ours as they consider aggregate-level data while we use individual-level data. Our paper also differs from these papers in that we consider matching with contracts. Finally, a number of papers estimate matching models with *transferrable* utility.[8] Among these, Akkus and Hortaçsu (2007) and Park (2011) are close to ours in that they study the merger of banks and mutual funds, respectively. Our paper adds to these papers by explicitly considering the effect of post-merger competition.

The second strand of the related literature is the large and growing theoretical literature on matching.[9] Our paper builds on Hatfield and Milgrom (2005, hereafter denoted as HM), who study a matching model with contract. In fact, our theoretical characterization is derived using the nested fixed-point algorithm, which contains HM's generalized Gale-Shapley algorithm as the inner loop. HM characterize the set of stable allocations for the matching model with contract. HM show that the set of stable allocations in two-sided matching with contracts has lattice property and proves the existence using Tarski's fixed-point theorem.[10] We follow their approach and incorporate two additional features; we allow externalities and participation decisions. In matching models, a player's individual rationality condition is about whether he has incentive to be matched with others, but not about his participation to the matching market. We explicitly consider the incentive to

[7] Although there is a transfer, our model is a model with non-transferrable utility. This is because players do not value transfers in the same way.

[8] See, e.g., Choo and Siow (2006), Fox (2010), Galichon and Salanié (2010), Bacarra *et al.* (2012), and Chiappori, Salanié, and Weiss (2010) .

[9] See, e.g., a survey by Roth (2008).

[10] See also Adachi (2000), Echenique and Oviedo (2006), and Ostrovsky (2008) for characterizations of the set of stable matchings using similar techniques in various matching environments.

"participate," and make the preference dependent on the number of "participants," which is the externalities we consider.

To the best of our knowledge, Sasaki and Toda (1996) and Hafalir (2008) are the only papers that investigate a two-sided matching model with externalities. Both papers consider a very general form of externalities. Analyzing such matching models is difficult because preference is defined over the set of assignments rather than matchings. Hence, regular definition of "stability" or "deviation" are not sufficient to analyze such a model because a deviating pair's preference also depends on how other agents would react to their deviation, not just their matching. To model how other agents would react to a player's deviation, both papers use what they call the *estimation function* approach. Estimation functions specify the expectations on the assignment (i.e., what the matching among all players would be) after each deviation. They prove that a strong requirement on the estimation function is necessary in order to guarantee the existence of stable matching. Based on their estimation function approach while considering a particular form of externalities (the payoff depends only on the number of operating firms in the market), we show the existence and provide characterizations.

The third strand of the literature is the literature on estimating entry models following Bresnahan and Reiss (1991) and Berry (1992), where the firms' underlying profit functions are inferred from the observed entry decisions.[11] We add to this literature by examining the entry and merger decisions jointly. We do so by combining two-sided matching model with the entry model. Our approach is similar to Ciliberto and Tamer (2009) in using a set estimator to address multiplicity of equilibria. Also, our identification argument builds on their identification results. Complementary to our study, Perez-Saiz (2012) considers a similar question and adopts an extensive form game for the merger and entry process in the U.S. cement industry. He models merger decisions to be conditional on entry decisions, while these decisions are joint decisions in our model. Other related studies include Jia (2008) and Nishida (2012) that characterize the equilibrium of an entry model with correlated markets as a fixed point in lattice and solve for it to estimate the model. Our paper differs from theirs in that ours make no equilibrium selection assumption and construct moment inequalities exploiting the lattice properties.

The fourth strand is the literature on horizontal merger decisions. In spite of the large literature considering the effects of mergers, studies on the endogenous horizontal merger decision itself are limited.[12] Kamien and Zang (1990) show limits to monopolization through mergers, and Qiu and Zhou (2007) point out the importance of firm heterogeneity in horizontal mergers. Gowrisankaran (1999) develops a computable dynamic industry competition

[11] Other recent contributions include Mazzeo (2002) and Seim (2006).

[12] There are also a few papers that study the relationship between merger decisions and merger review policy. See, e.g., Nocke and Whinston (2010).

model with an endogenous merger decision. Pesendorfer (2005) finds the relationship between market concentration and the profitability of mergers using a repeated game with merger decision. In a model with de novo foreign direct investment and cross-border merger and acquisitions, Nocke and Yeaple (2007) show the importance of firm heterogeneity as key determinants. Our paper adds to this literature by empirically investigating the role of firm heterogeneity in merger decisions.

Finally, our paper is also related to the literature studying banks' branching decisions. Ruffer and Holcomb (2001) use data from California and investigate the determinants of a bank's expansion decision by building a new branch and acquiring an existing branch, respectively. Their results show that a large bank would be likely to enter a new market by acquisition, but not through building a new branch, which is consistent with our result that larger potential entrants have higher synergy ceteris paribus. Wheelock and Wilson (2000) study the determinants of bank failures and acquisitions using a competing-risks model. Consistent with our finding, they show that less capitalized banks are more likely to be acquired for the period between 1984 and 1993. Cohen and Mazzeo (2007) estimate an entry model with vertical differentiation among retail depository institutions, and find evidence of product differentiation depending on market geography.

2 Model

We model the entry and merger decisions as a two-sided matching problem with externalities. We build the model combining models of entry (Bresnahan and Reiss, 1991, and Berry, 1992) and two-sided matching with contracts (Hatfield and Milgrom, 2005). After describing the model, we propose the solution concept that addresses externalities, $\mathcal{N}$-stability. Then, we provide characterizations of the $\mathcal{N}$-stable outcomes: the set of $\mathcal{N}$-stable outcomes forms a complete lattice, and the two extremum points of the set can be obtained by running a deferred acceptance algorithm that we propose. We use these characterizations for identification and estimation of the model.

2.1 A Matching Model of Entry and Merger

We consider a static entry model in which firms can use merger and acquisition as a form of entry in addition to regular entry without merger. In particular we integrate entry and merger decisions into a two-sided matching model between an incumbent firm (denoted by $i \in \{1, ..., N_{\mathcal{I}}\} \equiv \mathcal{I}$) and a potential entrant (denoted by $e \in \{1, ..., N_{\mathcal{E}}\} \equiv \mathcal{E}$). The model is a static one and being an incumbent is simply a characteristic of a firm. In other words, being an incumbent does not have particular dynamic implications. Note that we abstract from mergers between incumbents as well as mergers involving more than three firms.

We adopt a two-sided matching model because mutual consent between the two parties is required for a merger. In particular, we consider one-to-one two-sided matching instead of coalition formation, one-to-many matching, or many-to-many matching for the reason that the vast majority of mergers in our data are one-to-one mergers between an incumbent and a potential entrant. This fact reflects the types of market we observe: our data is from small regional markets where average number of incumbents are less than 5 (see Section 3 for the detail), thus mergers between incumbents or mergers involving more than two incumbents tend to be infeasible due to antitrust concerns. Finally, we consider a matching model instead of a specific extensive form game because the details about individual merger process (such as how investment banks and FDIC are involved in each case, how both sides negotiate the merger contract, etc.) are not observable in general.

Our matching model adds two features to a regular two-sided one-to-one matching model with contracts by i) considering two outside options (the decisions of "enter without merger" and "not to enter"), and ii) incorporating externalities that depend on the number of operating firms (network externalities).

Firms on both sides have three types of choices. Potential entrant e can choose not to enter (denoted by $\{o\}$), enter by itself (denoted by $\{e\}$), or merge with incumbent i with a merger contract k_{ei} (denoted by $\{k_{ei}\}$). Merger contracts are bilateral ones between a potential entrant and an incumbent, and each firm can sign only one merger contract with a firm on the other side. A contract $k_{ei} = (e, i, p_{ei})$ specifies a potential entrant and an incumbent pair and the terms of the merger, $p_{ei} \in P$, where the set of merger terms P is finite as in HM.

Similarly to potential entrants, incumbent i has three types of choices: it can choose not to enter (denoted by $\{o\}$), enter by itself (denoted by $\{i\}$), or merge with potential entrant e with a merger contract k_{ei} (denoted by $\{k_{ei}\}$).

We denote the set of merger contracts by $\mathcal{K} \equiv \mathcal{E} \times \mathcal{I} \times P$. Note that the set of merger contracts does not include the case where a potential entrant or incumbent does not enter the market or the case where they enter by themselves. Let the set of merger contracts in which entrant e is involved be K_e and in which incumbent i is involved be K_i, i.e.,

$$K_e = \bigcup_{i \in \mathcal{I}} \{k_{ei}\}, \text{ and } K_i = \bigcup_{e \in \mathcal{E}} \{k_{ei}\}.$$

We also define the set of available choices for e and i as

$$\overline{K}_e = K_e \cup \{e\} \cup \{o\} \text{ and } \overline{K}_i = K_i \cup \{i\} \cup \{o\}.$$

For simplifying the notation, we define $\overline{\mathcal{K}} = \bigcup_{j \in \mathcal{E} \cup \mathcal{I}} \overline{K}_j$.

Payoffs depend on the outcome of the matching game. Because we consider entry and merger decisions after which firms compete, a firm's payoff is affected not only by its entry and merger decisions but also by other firms' entry and merger decisions. Hence, we need to consider a model with externalities due to post-entry competition. To address these externalities, we follow Bresnahan and Reiss (1991) and Berry (1992) and assume that the number of operating firms, N, negatively affects the payoff of the firm (note that N depends on entry and merger decisions jointly). Furthermore, as non-monetary components such as a manager's idiosyncratic taste over the potential merger may play an important role in the actual merger decisions, e.g., Malmendier and Tate, 2008), we allow the payoff of a firm to depend not only on the profit but also on other factors. We denote the payoff of incumbent i as

$$\begin{array}{ll} \Pi_i(k_{ei}, N) = \frac{1}{\sigma_i}\pi_i(k_{ei}, N) + \varepsilon_{ei} & \text{if firm } i \text{ merges with firm } e \text{ with contract } k_{ei}, \\ \Pi_i(i, N) = \frac{1}{\sigma_i}\pi_i(i, N) + \varepsilon_{ii} & \text{if firm } i \text{ enters without merger,}(se \\ \Pi_i(o, N) = 0 & \text{if firm } i \text{ does not enter,} \end{array}$$

where $\pi_i(\cdot, N)$ is the monetary profit with N operating firms, ε_{ei} and ε_{ii} are idiosyncratic shocks, and σ_i is a parameter that captures the relative importance of idiosyncratic shocks to the monetary profit. Reflecting the negative externalities of the post-entry competition as in the empirical entry literature, we assume $\Pi_i(k, N) < \Pi_i(k, N-1)$ for all $k \in K_i \cup \{i\}$. Also, as we assume ε_{ei} and ε_{ee} to be continuously distributed, the preferences are strict generically.

In the same way, we can write potential entrant e's payoff as

$$\begin{array}{ll} \Pi_e(k_{ei}, N) = \frac{1}{\sigma_e}\pi_e(k_{ei}, N) + \varepsilon_{ie} & \text{if firm } e \text{ merges with firm } i \text{ with contract } k_{ei}, \\ \Pi_e(e, N) = \frac{1}{\sigma_e}\pi_e(e, N) + \varepsilon_{ee} & \text{if firm } e \text{ enters without merger,} \\ \Pi_e(o, N) = 0 & \text{if firm } e \text{ does not enter,} \end{array}$$

where $\pi_e(\cdot, N)$ is the monetary profit with N operating firms, ε_{ei} ($\neq \varepsilon_{ie}$) and ε_{ee} are idiosyncratic shocks, and σ_e is a parameter that captures the relative importance of idiosyncratic shocks. We do not impose $\varepsilon_{ei} = \varepsilon_{ie}$ since acquiring and target firms may have heterogenous preferences on the potential merger. As in the case of the incumbents, we assume $\Pi_e(k, N) < \Pi_e(k, N-1)$ for all $k \in K_e \cup \{e\}$, and preferences are strict generically.

Note that the model is a matching model with non-transferable utility, though the terms of the merger contract may include transfers between the firms. This is because we allow σ_i and σ_e to be different across firms, reflecting the fact that factors not directly measured by monetary profit may have varying importance across firms in merger and entry decisions. Such differences may result from variations in the degree to which managerial

and shareholder interests are misaligned as in Jensen and Meckling (1976). Other sources could be differences in CEOs' overconfidence on merger decisions (Malmendier and Tate, 2008) and variations in manager's strategic ability on entry decisions (Goldfarb and Xiao, 2011).

2.2 Solution Concept: $\mathcal{N}$-Stable Outcome

2.2.1 Estimation Function and Chosen Set

We extend the solution concept of stable allocation of Hatfield and Milgrom (2005) to our environment with externalities. Considering externalities in two-sided matching is difficult in general (Sasaki and Toda, 1996, Hafalir, 2008). This is because preference is dependent not only on one's own matching (as in the case without externalities), but also on how other players are matched with each other (i.e., the entire assignment). This dependence of preference on the entire assignment requires us to consider how players expect the reaction of other players in thinking about the definition of stability that is appropriate for the matching model with externalities.

In order to describe the way each player expects how other players are matched with each other, we follow Sasaki and Toda (1996) and Hafalir (2008) to use the *estimation function* approach. In a matching model with externalities, if a pair of players deviates (dissolves the match), each member of the pair has to think not only about their own matching but also about how other players (including his/her previous partner) are reacting to the deviation because preferences are defined over assignment rather than matching.[13] Sasaki and Toda (1996) and Hafalir (2008) define the estimation function as a mapping from the set of possible matches to the set of matchings of all players, which specifies the expected assignments resulting from a deviation. In our notation, the estimation function of firm j, denoted by $\mathcal{N}_j$, is $\mathcal{N}_j : \overline{K}_j \to \overline{\mathcal{K}}$ if we consider a general form of externalities.

In our environment, however, we focus on the (network) externalities that enter into the firm's payoff only through the number of operating firms in the market, N (note that this depends on the outcome of matchings of other firms as well).[14] Thus, we define the estimation function for firm j as $\mathcal{N}_j : \overline{K}_j \to \mathbb{R}_+$, a mapping from the possible choices to

[13] To illustrate this point, suppose Players A and X are currently matched. If Player A deviates to form a blocking pair with Player Y, Player A has to consider not only that Player Y has incentive to be matched with A, but also what other players, including Player X, would do after the deviation, because the entire outcome (rather than just whom Player A is matched with) affects Player A. Thus, Player A's expectations about the possible entire outcomes after the deviation are crucial.

[14] This assumption can be relaxed in some dimensions such as the case that the firms are vertically differentiated as in Mazzeo (2002). Suppose each firm has its *type* based on some observable firm specific characteristics such as high, medium, and low quality or for- and non-profit. Denote the number of total operating firms for each type by $N_s, s = 1, 2, ..., S$. Then, the payoff can depend on the total number of operating firms for each *type*: $\pi_j(\{k\}, (N_s)_{s=1}^S)$, and similar analysis can be extended.

the *estimated* number of operating firms.[15] For example, $\mathcal{N}_e(e) = 3$ denotes that potential entrant e estimates the number of operating firms to be 3 if e enters the market without merger, and the function $\mathcal{N}_e(\cdot)$ specifies the estimated number of operating firms for all other possible choices o and k_{ei} as well, such as $\mathcal{N}_e(o) = 2$ and $\mathcal{N}_e(k_{ei}) = 5$. This restriction on the form of externalities allows analysis and characterization to be more tractable than those of existing works.

Now we describe the choice by firms given the estimation function. In order to represent the choice by a potential entrant given the set of available merger contracts K_e and the estimation function $\mathcal{N}_e$, we define the *chosen set from merger contracts* $C_e(K_e, \mathcal{N}_e)$ for e as the following:

$$C_e(K_e, \mathcal{N}_e) = \begin{cases} \varnothing \text{ if } \max\{\Pi_e(o, \mathcal{N}_e), \Pi_e(e, \mathcal{N}_e)\} \geq \max_{k \in K_e}\{\Pi_e(k, \mathcal{N}_e)\} \\ k_{ei} \text{ if } \Pi_e(k_{ei}, \mathcal{N}_e) \geq \Pi_e(k, \mathcal{N}_e), \ \forall k \in \overline{K}_e, \end{cases}$$

where $\Pi_e(k, \mathcal{N}_e) \equiv \Pi_e(k, \mathcal{N}_e(k))$ by suppressing k from $\mathcal{N}_e(k)$. This set is the best available merger contract given the available set of contracts K_e, which can be a null set if no merger contract is more attractive than de novo entry and no entry. Similarly, we define the chosen set from merger contracts for incumbent i given the available set of contracts K_i as follows:

$$C_i(K_i, \mathcal{N}_i) = \begin{cases} \varnothing \text{ if } \max\{\Pi_i(o, \mathcal{N}_i), \Pi_i(i, \mathcal{N}_i)\} \geq \max_{k \in K_i}\{\Pi_i(k, \mathcal{N}_i)\} \\ k_{ei} \text{ if } \Pi_i(k_{ei}, \mathcal{N}_i) \geq \Pi_i(k, \mathcal{N}_i), \ \forall k \in \overline{K}_i, \end{cases}$$

which also can be a null set if no merger contract is more attractive than de novo entry and no entry.

In addition to the chosen set from merger contracts, we also need to track the choice of firms including no entry and entry without merger. We denote the *chosen set* for incumbent i by $\overline{C}_i(\overline{K}_i, \mathcal{N}_i)$ and for potential entrant e by $\overline{C}_e(\overline{K}_e, \mathcal{N}_e)$, which is either the most preferred merger contract available to firm j in K_j, entry without merger, or no entry, i.e.,

$$\begin{aligned} \overline{C}_i(\overline{K}_i, \mathcal{N}_i) &= \operatorname{argmax}_{k \in \overline{K}_i}\{\Pi_i(k, \mathcal{N}_i)\}, \\ \overline{C}_e(\overline{K}_e, \mathcal{N}_e) &= \operatorname{argmax}_{k \in \overline{K}_e}\{\Pi_e(k, \mathcal{N}_e)\}. \end{aligned}$$

The difference between the *chosen set from merger contracts* and the *chosen set* is that $C_i(K, \mathcal{N}_i)$ and $C_e(K, \mathcal{N}_e)$ take the null set if any merger contract in K is not preferred to no entry or entry without merger, while $\overline{C}_i(\overline{K}_i, \mathcal{N}_i)$ and $\overline{C}_e(\overline{K}_e, \mathcal{N}_e)$ specify the optimal choice among $\overline{K}_j$ for each player, which may include no entry or entry without merger.

[15] Note that we allow the estimated number of operating firms not only to be a non-negative integer, but also a positive real number. This is because we allow mixing in the case of indifference as we will discuss in Section 2.2.2.

2.2.2 $\mathcal{N}$-Stability

In matching models without externalities, a matching is *stable* if it satisfies i) the no-blocking-pair condition and ii) individual rationality. The no-blocking-pair condition requires that there be no pair of players who are weakly better off than they would be with their current match. Individual rationality requires that no player form a couple that is less preferable than being unmatched. Our version of stability adds to the standard definition of stability in two ways. First, to address the issue resulting from the presence of externalities, we modify the solution concept of φ-stability in Sasaki and Toda (1996) and Hafalir (2008).[16] Second, we slightly modify the individual rationality so that firms can choose to be unmatched in two ways: firms can choose "not to enter the market" ($\{o\}$) or "to enter the market without merger" ($\{i\}$ and $\{e\}$).

Externalities require an additional condition on the regular solution concept of stability because the players have an *estimation function* about what might happen after each choice, and the behavior of players needs to be consistent with the estimation function. Sasaki and Toda (1996) and Hafalir (2008) call this condition φ-admissibility. They use this condition to define their solution concept for matching models with externalities (called φ-stability). We follow their approach and consider a solution concept that incorporates our estimation function $\mathcal{N}$ discussed above (our estimation function is simpler than theirs), which we call $\mathcal{N}$-stability.

The way we incorporate the estimation function $\mathcal{N}$ to our solution concept is by requiring the "consistency" of the estimation function, i.e., each firm's *estimation* on the number of operating firms equals the actual number of operating firms. In order to define consistency, let us first define correspondence λ, which maps from the set of all the firms' estimation functions and the set of available contracts to positive real numbers. This correspondence, λ, computes the resulting number of operating firms given the chosen set of all players and estimation function $\mathcal{N} = \{\mathcal{N}_j\}_{j\in\mathcal{E}\cup\mathcal{I}}$, i.e.,

$$\lambda(\overline{K},\mathcal{N}) = \frac{1}{2}\left[N_{\mathcal{E}} + N_{\mathcal{I}} - \sum_{j\in\mathcal{E}\cup\mathcal{I}} p_j(o;\overline{K},\mathcal{N}_j) + \sum_{j\in\mathcal{E}\cup\mathcal{I}} p_j(j;\overline{K},\mathcal{N}_j)\right],$$

where $p_j(k;\overline{K}_j,\mathcal{N}_j)$ is a probability distribution over the set $\overline{C}_j(\overline{K}_j,\mathcal{N}_j)$ defined as follows:

$$p_j(k;\overline{K}_j,\mathcal{N}_j) = \begin{cases} 1 & \text{if } \{k\} = \overline{C}_j(\overline{K}_j,\mathcal{N}_j) \\ q_k & \text{if } k \in \overline{C}_j(\overline{K}_j,\mathcal{N}_j) \text{ and } \{k\} \neq \overline{C}_j(\overline{K}_j,\mathcal{N}_j) \\ 0 & \text{if } k \notin \overline{C}_j(\overline{K}_j,\mathcal{N}_j) \end{cases},$$

[16] The concept of φ-stability extends the regular concept of stability with an additional condition regarding the estimation function φ, which addresses the issue of externalities.

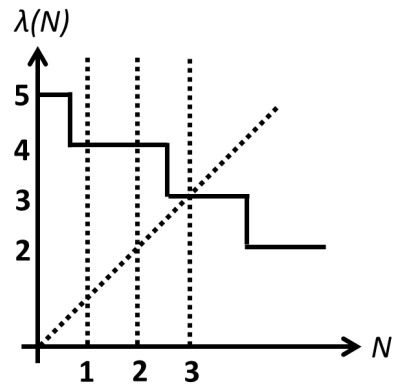

Figure 1: An example of $\lambda(\overline{K}, N)$ for a given $\overline{K}$ (we suppress the dependence on $\overline{K}$). N is estimation on number of entering firms, and $\lambda(\overline{K}, N)$ is the resulting number of entering firms given estimation N. We show that $\lambda(\overline{K}, N)$ is decreasing in N in Lemma 1. 45-degree line corresponds to consistency of estimation.

where $\sum_{\kappa \in \overline{C}_j(\overline{K}_j, \mathcal{N}_j)} q_\kappa = 1$ with $q_\kappa \in [0, 1]$. The correspondence λ simply returns the number of operating firms. Figure 1 presents an example of λ. Note that we use $p_j(o; \overline{K}, \mathcal{N}_j)$ instead of $1\left\{\overline{C}_j(\overline{K}, \mathcal{N}_j) = o\right\}$ to allow for mixing of choices by the players in a nongeneric case when $\overline{C}_j(\overline{K}_j, \mathcal{N}_j)$ is not a singleton (i.e., $\{k\} \neq \overline{C}_j(\overline{K}_j, \mathcal{N}_j)$). Such non-generic case may occur if a firm is indifferent between alternatives. For a generic case that the chosen set is singleton (i.e., $\{k\} = \overline{C}_j(\overline{K}_j, \mathcal{N}_j)$), the choice probability for the alternative is 1. As the correspondence λ aggregates choice probabilities of each firm, these imply that the range of $\lambda(\cdot, \cdot)$ is $\mathbb{R}_+$.

We allow mixing of choices by a player in a nongeneric case of indifference in order to guarantee the existence of a solution.[17] Indifference occurs, for example, in a nongeneric case where a firm's profit from entering alone happens to be zero. Such cases occur for some specific values of N, corresponding to the vertical jump of $\lambda(\overline{K}, N)$ in Figure 1 where $\lambda(\overline{K}, N)$ takes a set value between two integers. In such a case, we consider that the firm may mix between the indifferent choice alternatives with probability q_k for $k \in \overline{C}_j(\overline{K}_j, \mathcal{N}_j)$. This implies that λ is in fact a correspondence (because q_k can take any value with $q_k \in [0, 1]$ and $\sum_{\kappa \in \overline{C}_j(\overline{K}_j, \mathcal{N}_j)} q_\kappa = 1$), which we later use to apply Kakutani's Fixed Point Theorem to guarantee existence of the solution as discussed further in Section 3.2.

[17] Note that the discreteness is one of the reasons why guaranteeing existence of stable matching is difficult in general when externalities are present. In our case, we approach this difficutly by relying on the fact the indifference occurs in a particular way as described (and thus we allow for mixing), so that we can show existence and characterize the stable outcome.

Also, for the case of indifference, it is known that stability does not imply Pareto efficiency. For our purpose, we require existence of stable outcome, but not Pareto efficiency.

Now we can write the consistency requirement: the set of available choices $\overline{K}$ and estimation function $\mathcal{N}$ are such that the number of operating firms equals the *estimated* number of firms, i.e.,

Condition 1 (Consistency of Estimation) *A set of available choices and estimation function* $(\overline{K}, \mathcal{N})$ *is consistent if*

$$\mathcal{N}_j(k_j) \in \lambda(\overline{K}, \mathcal{N}), \quad \forall j \in \mathcal{E} \cup \mathcal{I},$$

where $k_j \in \overline{C}_j(\overline{K}, \mathcal{N}_j)$.

In Figure 1, this condition corresponds to the point on 45-degree line where λ intersects. This is because the estimated number of operating firm equals the resulting number of operating firms.

The second modification to the standard definition of stability is that we allow two outside options: entry without merger and no entry. In the standard definition of individual rationality for matching models, players compare being unmatched to matching with someone. In such cases players can unilaterally choose to stay in the market alone. In our case, the players choose one of the better outside options if they were not matched, and they can make this decision *unilaterally*. We write this condition as follows [18]:

Condition 2 (Individual Rationality) *A set of available choices and estimation functions* $(\overline{K}, \mathcal{N})$ *is individually rational if* $\forall j \in \mathcal{E} \cup \mathcal{I}, \nexists \widetilde{k} \in \overline{K}_j$ *s.t.*

$$\Pi_j(\widetilde{k}, \mathcal{N}_j) \geq \Pi_j(\overline{C}_j(\overline{K}, \mathcal{N}_j), \mathcal{N}_j).$$

The last condition we require is the *no-blocking-contract* condition. It requires that there exist no merger contracts to which firms from both sides would be willing to deviate. This condition is a standard one to define stability in a two-sided matching model with contracts (see, e.g., HM). To define this condition, let us first define the set of merger contracts included in the set of outcomes $\overline{K}$ by K, i.e., $K = \bigcup_{j \in \mathcal{E} \cup \mathcal{I}} \overline{K}_j \backslash (\{o\} \cup \{j\})$. This is simply the set of merger contracts chosen given $\overline{K}$. Using this notation, the no-blocking-contract condition can be written as follows:

Condition 3 (No Blocking Contracts) *A set of available choices and estimation functions* $(\overline{K}, \mathcal{N})$ *admits no blocking contracts if* $\nexists \widetilde{k} \subset \mathcal{K}$ *s.t.* $\widetilde{k} \neq K$ *and*

$$\widetilde{k} = \bigcup_{i \in \mathcal{I}} C_i(K \cup \widetilde{k}, \mathcal{N}_i) = \bigcup_{e \in \mathcal{E}} C_e(K \cup \widetilde{k}, \mathcal{N}_e).$$

[18] We slightly abuse notation when the chosen set $\overline{C}_j(\overline{K}, \mathcal{N}_j)$ is not a singleton. In such cases, the payoff to firm j is exactly the same regardless of the choices in $\overline{C}_j(\overline{K}, \mathcal{N}_j)$.

Now we define our solution concept of $\mathcal{N}$-*stabile outcome* using the three conditions defined above.

Definition 1 ($\mathcal{N}$-stable outcome) *A set of available choices and estimation functions* $(\overline{K}^*, \mathcal{N}^*)$ *is a* $\mathcal{N}$*-stable outcome if Conditions 1,2, and 3 are satisfied.*

In the next section, we propose a deferred acceptance algorithm, which accommodates externalities in order to show the existence of $\mathcal{N}$-stable outcomes, and provide characterizations. We use those characterizations as the basis of our estimation strategy.

2.3 A Generalized Gale-Shapley Algorithm with Externalities

In this section, we show the existence of $\mathcal{N}$-stable outcome by proposing a generalized Gale-Shapley algorithm with externalities, and also show some of the properties of the $\mathcal{N}$-stable outcome, which we use for our identification and estimation. The algorithm we propose uses HM's generalized Gale-Shapley algorithm as an inner loop conditional on the estimated number of firms. The outer-loop adjusts firms' estimation in order to find the estimated number of operating firms, N^*, such that it equals the actual number of operating firms, i.e., satisfying the Consistency of Estimation, $N^* = \mathcal{N}_j(k_j^*) \in \lambda(\overline{K}^*, \mathcal{N}^*)$ (Condition 1).

First, we consider an estimation function $\mathcal{N}$ that is part of the definition of the $\mathcal{N}$-stable outcome. We use the following estimation function: $\forall j \in \mathcal{E} \cup \mathcal{I}$, $\forall k \in \overline{K_j}$,

$$\mathcal{N}_j(k) = N,$$

for some $N < N_{\mathcal{E}} + N_{\mathcal{I}}$.[19] This describes that the number of estimated operating firms remains unchanged by firm j's choice k. In other words, there are certain number of firms that the market can accommodate, and firm j's choice does not affect this number. Note that this estimation function does not exclude the case of other firms changing their behavior.[20]

[19] Note that we do not allow $N = N_{\mathcal{E}} + N_{\mathcal{I}}$ for the reason that there always exist some potential entrants that do not enter in our data.

[20] Baccara et al.(2012) also consider matching with externalities, and they assume that player's behavior does not change after any deviation. Using estimation function approach, their assumption in our context can be described as follows:

$$\mathcal{N}_j(k) = \begin{cases} N \text{ if } k = j \\ N-1 \text{ if } k = k_{jl} \\ N-1 \text{ if } k = o. \end{cases}$$

Note that this estimation function does not necessarily imply their assumption (while their assumption implies this estimation function). This is because the above estimation function only concerns the *aggregate* number of operating firms; hence the same firm may not necessarily behave the same way after different

One property related to N plays an important role for us to show existence through our algorithm: the number of operating firms given estimation $\mathcal{N}$ is weakly decreasing in N. This property is directly driven by the property that the payoff of the firms negatively depends on N. Because λ is a correspondence, we define function $\lambda^{\min}(\overline{K}, \mathcal{N}) \equiv \min \lambda(\overline{K}, \mathcal{N})$ and $\lambda^{\max}(\overline{K}, \mathcal{N}) \equiv \max \lambda(\overline{K}, \mathcal{N})$, which are simply functions that take the minimum and maximum values of the correspondence $\lambda(\overline{K}, \mathcal{N})$, to illustrate this property.

Lemma 1 *Given the estimation function $\mathcal{N}_j(k) = N$, $\forall k \in \overline{K}$, the functions $\lambda^{\min}(\overline{K}, \mathcal{N})$ and $\lambda^{\max}(\overline{K}, \mathcal{N})$ are weakly decreasing in N. Also, $\lambda(\overline{K}, \mathcal{N})$ is upper-hemicontinuous.*

All the proofs are in Appendix A. This Lemma implies that as the estimation becomes optimistic (with a small number of operating firms), the number of firms that actually enter increases, and vice versa. This property, combined with HM's generalized Gale-Shapley algorithm, yields a fixed point for N^* using the algorithm we propose below.

Our algorithm is a nested fixed-point algorithm that consists of inner and outer loops. The inner loop is a variation of HM's generalized Gale-Shapley algorithm conditional on an estimated number of firms, and the outer loop is a simple minimization algorithm that runs to find a fixed point corresponding to the consistency of estimation (Condition 1).

A Generalized Gale-Shapley Algorithm with Externalities

1. (Outer Loop) Set $N = 0$ as an initial value for N. Use any minimization routine[21] to minimize $\left\|\lambda^{\min}(\overline{K}, N) - N\right\|$.

2. (Inner Loop) Run the following ($\mathcal{E}$-proposing) generalized Gale-Shapley algorithm given N:

   (a) Initialize $K_{\mathcal{E}} = \mathcal{K}$, $K_{\mathcal{I}} = \varnothing$.

   (b) All $e \in \mathcal{E}$ choose $\{e\}, \{o\}$, or make the offer that is most favorable to e from $K_{\mathcal{E}}$ to members of $\mathcal{I}$.

   (c) All $i \in \mathcal{I}$ consider $\{i\}, \{o\}$, and all available offers, then hold the best, and reject the others.

   (d) Update $K_{\mathcal{E}}$ by removing offers that have been rejected. Update $K_{\mathcal{I}}$ by including newly made offers.

deviations. We provide the existence result and an algorithm to find a stable outcome for this estimation function in the Supplementary Material. The idea is to combine the generalized Gale-Shapley algorithm with externalities we propose and the EM algorithm for the distribution of players with different types of estimations.

[21] Because $\left\|\lambda^{\min}(\overline{K}, N) - N\right\|$ is quasiconvex in N, global minima can easily be obtained by any one-dimensional optimization algorithm such as the Bracketing Algorithm, Newton's method, or grid search.

(e) If there is no change to $K_{\mathcal{E}}$ and $K_{\mathcal{I}}$, count the number of operating firms $\lambda^{\min}(K_{\mathcal{E}} \cap K_{\mathcal{I}}, N)$ and go to Step 3. Otherwise, return to Step (b).

3. (Outer Loop) If the convergence criterion of the minimization routine is not satisfied, obtain N for the next round from the minimization routine, and go to step 2. If the convergence criterion of the minimization is satisfied, terminate the algorithm.

We can also consider $\mathcal{I}$-proposing algorithm, which would entail making the following changes to Step 2 above: substituting $K_{\mathcal{I}} = \mathcal{K}$, $K_{\mathcal{E}} = \varnothing$ with $K_{\mathcal{E}} = \mathcal{K}$, $K_{\mathcal{I}} = \varnothing$ in (a); e with i in (b); i with e in (c); and $K_{\mathcal{E}}$ with $K_{\mathcal{I}}$ and $K_{\mathcal{I}}$ with $K_{\mathcal{E}}$ in (d).

The inner loop of our algorithm corresponds to HM's generalized Gale-Shapley algorithm because the estimation of N is fixed. In each round each potential entrant offers a merger contract to an incumbent or chooses either of the outside options in step (b), and each incumbent holds the best contract offered or either of the outside options in step (c). The set of available contracts for $\mathcal{I}$, $K_{\mathcal{I}}$, starts with an empty set and expands *monotonically* as more offers are (cumulatively) made each round. The set of available contracts for $\mathcal{E}$, $K_{\mathcal{E}}$, starts with the entire set of contracts in step (a), and it *monotonically* shrinks as offers are rejected each round. HM show the existence of stable allocation and the characterization using this (inner-loop) algorithm as follows.

Theorem 1 (Hatfield and Milgrom (2005)) *(i) The $\mathcal{E}$-proposing inner loop converges to the $\mathcal{E}$-optimal stable allocation given N, $\overline{K}^{*\mathcal{E}}(N)$, and the $\mathcal{I}$-proposing inner loop converges to the $\mathcal{I}$-optimal stable allocation given N, $\overline{K}^{*\mathcal{I}}(N)$;*

(ii) The set of stable contracts $\overline{K}^{*\mathcal{E}}(N)$ is unanimously the most preferred set of contracts among all of stable contracts for $\mathcal{E}$ and unanimously the least preferred for $\mathcal{I}$, and vice versa for $\overline{K}^{*\mathcal{I}}(N)$;

(iii) The number of operating firms $\lambda(\overline{K}, N)$ is the same for all stable allocations given N;

(iv) An unmatched firm in a stable allocation is also unmatched in any stable allocation given N.

Parts (iii) and (iv) of Theorem 1 are the so-called "rural hospitals theorem" (Roth, 1986). In our case, this indicates that the set of firms that are unmatched is the same for all stable outcomes,[22] which in turn implies that the number of operating firms, $\lambda(\overline{K}, \mathcal{N})$, is

[22] Hatfield and Milgrom (2005) show a variation of the rural-hospitals theorem in the case of many-to-one matching with contract. They show that every hospital signs exactly the same number of contracts at every point in the set of stable allocations if the hospitals' preferences satisfy what they call the law of aggregate demand and substitutability. This result implies that the set of hospitals that cannot fill the capacity in a stable allocation cannot fill it in any stable allocation. It does not necessarily imply the same set of doctors is hired in all stable allocations, though the number of unmatched doctors remains the same.

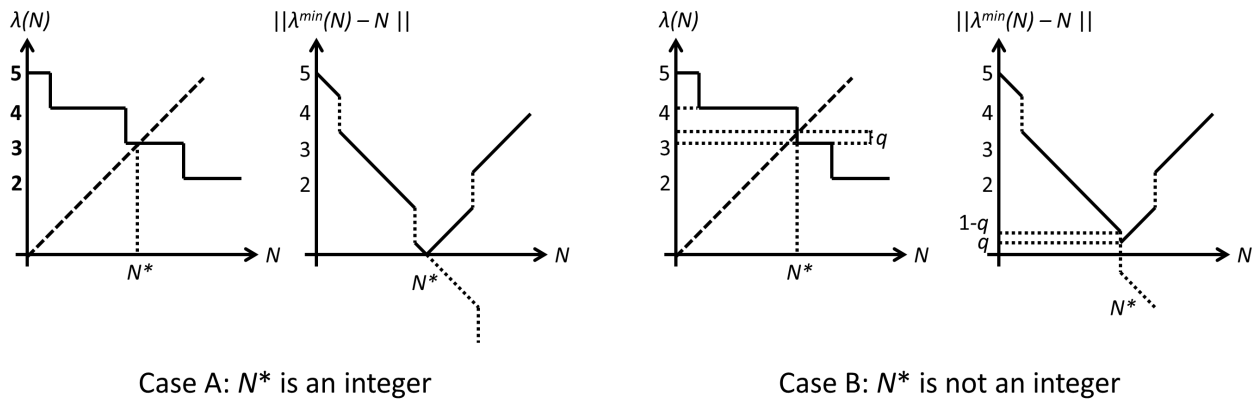

Figure 2: Finding consistent N. The two pictures on the left correspond to the case where N^* is an integer. The two pictures on the right correspond to the case where there is mixing, with probability q. The function $\left\|\lambda^{\min}(\overline{K}, N) - N\right\|$ is quasi-concave, and converges globally to N^* by applying any regular minimization alogorithm. In the figure, we suppress the dependence of $\lambda^{\min}$ on $\overline{K}$.

the same for all stable outcomes given N.[23] Thus, $\lambda(\overline{K}, \mathcal{N})$ is the same in both $\mathcal{E}$-proposing and $\mathcal{I}$-proposing algorithms.

Given this property of the inner loop, the outer loop of our algorithm finds the estimated number of operating firms N satisfying the consistency condition, $N \in \lambda(\overline{K}, \mathcal{N})$. The following proposition shows there exists such N.

Proposition 1 *There exists N^*such that $N^* \in \lambda(\overline{K}^*, \mathcal{N}^*)$.*

Now, using Theorem 1 together with Proposition 1, we show that our algorithm converges to a stable outcome.

Theorem 2 *The generalized Gale-Shapley algorithm with externalities globally converges to an $\mathcal{N}$-stable outcome. If the algorithm starts from $(K_{\mathcal{E}}, K_{\mathcal{I}}) = (\mathcal{K}, \varnothing)$, then it converges to the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome, $\overline{K}^{*\mathcal{E}}$. Similarly, if the algorithm starts from $(K_{\mathcal{E}}, K_{\mathcal{I}}) = (\varnothing, \mathcal{K})$, then it converges to the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome, $\overline{K}^{*\mathcal{I}}$.*

However, in our case of one-to-one matching with contract, preferences of both sides satisfy these two conditions. Therefore, we can show that the set of unmatched firms is identical in any stable outcome given N.

[23] Note that we need one more step to show the number of operating firms $\lambda(\overline{K}, N)$ is the same for all stable allocations given N because there are two types of outside options in our model. Consider a firm that chooses one of the outside options, say $j \in \{j, o\}$, in a $\mathcal{N}$-stable outcome. This implies that $\Pi_j(j, N)$ is higher than the payoff from not entering and the payoff from any available contract under *any* $\mathcal{N}$-stable outcome. Hence, the firm's chosen set under any $\mathcal{N}$-stable outcome remains to be $\{j\}$. Thus, the number of operating firms $\lambda(\overline{K}, N)$ is the same for all $\mathcal{N}$-stable outcomes given N.

Figure 2 describes the global convergence of $\left\|\lambda^{\min}(\overline{K}, N) - N\right\|$, which assures the convergence of the generalized Gale-Shapley algorithm with externalities. Monotonicity of λ in N implies that the function $\left\|\lambda^{\min}(\overline{K}, N) - N\right\|$ monotonically decreases below N^* and it monotonically increases above N^*, and the function is quasi-convex. In Case A where N^* is an integer, the algorithm stops when $\left\|\lambda^{\min}(\overline{K}, N) - N\right\|$ is minimized at N^* with value of 0. In Case B where $N^*$is not an integer, the algorithm stops when $\left\|\lambda^{\min}(\overline{K}, N) - N\right\|$ is minimized at N^* with value of q, which corresponds to the probability of the indifferent firm entering the market.

The following two corollaries are useful to characterize an $\mathcal{N}$-stable outcome in terms of payoffs and the number of operating firms, which give us the basis of the inference discussed in the estimation section. Let us first define the number of mergers given an allocation $\overline{K}$ as $\gamma(\overline{K}, \mathcal{N})$.

Corollary 1 *(i) The $\mathcal{E}$-proposing and $\mathcal{I}$-proposing generalized Gale-Shapley algorithms with externalities terminate at the same N^*.*

(ii) An unmatched firm j in an $\mathcal{N}$-stable outcome is also unmatched in any $\mathcal{N}$-stable outcome, and the payoff of firm j is the same in any $\mathcal{N}$-stable outcome.

(iii) In any $\mathcal{N}$-stable outcome $(\overline{K}^*, \mathcal{N}^*)$, the number of mergers is uniquely determined, i.e., $N^*_{merge} = \gamma(\overline{K}^*, \mathcal{N}^*)$ $\forall(\overline{K}^*, \mathcal{N}^*)$.

Parts (i) and (ii) of Corollary 1 state that the number of operating firms does not depend on the selection of equilibrium, and is uniquely determined. Part (ii) also implies that the payoff for an unmatched firm is also uniquely determined. Hence, we can obtain the payoff of unmatched firms in any $\mathcal{N}$-stable outcome using the $\mathcal{E}$-optimal and $\mathcal{I}$-optimal $\mathcal{N}$-stable outcomes, i.e., for any firm j choosing not to enter or entry without merger,[24]

$$\Pi_j(\overline{K}^*, \mathcal{N}^*) = \Pi_j(\overline{K}^{*\mathcal{E}}, \mathcal{N}^*) = \Pi_j(\overline{K}^{*\mathcal{I}}, \mathcal{N}^*). \tag{1}$$

Also, part (iii) of Corollary 1 implies the following equality for the number of mergers,

$$\gamma(\overline{K}^*, \mathcal{N}^*) = \gamma(\overline{K}^{*\mathcal{E}}, \mathcal{N}^*) = \gamma(\overline{K}^{*\mathcal{I}}, \mathcal{N}^*). \tag{2}$$

We use these equalities to construct moment equalities in our estimation.

Corollary 2 *The $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome is preferred to any other stable outcome by all $e \in \mathcal{E}$ and the least preferred by all $i \in \mathcal{I}$, i.e., for any $\mathcal{N}$-stable outcome $\overline{K}^*$,*

$$\Pi_e(\overline{K}^{*\mathcal{E}}, \mathcal{N}^*) \geq \Pi_e(\overline{K}^*, \mathcal{N}^*) \text{ and } \Pi_i(\overline{K}^*, \mathcal{N}^*) \leq \Pi_i(\overline{K}^{*\mathcal{E}}, \mathcal{N}^*) \tag{3}$$

[24] We use $\Pi_j(\overline{K}^*, \mathcal{N}^*)$ to denote $\Pi_j(k_j, \mathcal{N}^*)$ with $k_j \in \overline{K}^*$ for notational simplicity.

for the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome.

Similarly, the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome is preferred to any other stable outcome by all $i \in \mathcal{I}$ and the least preferred by all $e \in \mathcal{E}$, i.e., for any $\mathcal{N}$-stable outcome $\overline{K}^*$,

$$\Pi_i(\overline{K}^{*\mathcal{I}}, \mathcal{N}^*) \geq \Pi_i(\overline{K}^*, \mathcal{N}^*) \text{ and } \Pi_e(\overline{K}^{*\mathcal{I}}, \mathcal{N}^*) \leq \Pi_e(\overline{K}^*, \mathcal{N}^*) \tag{4}$$

for the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome.

This corollary generalizes part (ii) of Theorem 1 to our environment, and states that any equilibrium payoffs for incumbents are bounded above by the payoff in the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome and below by the payoff in the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome, and that any equilibrium payoffs for potential entrants are bounded above by the payoff in the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome and below by the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome. We use these inequalities to construct moment inequalities in our estimation.

3 Data

Before presenting the estimation and identification, let us discuss the data we use in the estimation. The banking industry in the U.S. provides us with an interesting and important change in federal regulation. The Riegle-Neal Interstate Banking and Branching Act of 1994 (the Act), enacted in 1997, permitted banks to establish branches nationwide by eliminating all barriers to interstate banking at the state level. Another regulation this legislation eliminated, which is more important for this study, is the regulation on *intrastate* branching, and more specifically the regulation on de novo entry for intrastate branching. Before this legislation went into effect, 13 states prohibited intrastate de novo branching and permitted branching only by merger, while the other 37 states fully permitted intrastate branching.

We use the data on commercial banks in the U.S. from the local markets of these 13 states in which intrastate branching became fully permitted by the Act, which allows us to study the effect of the intrastate branching regulation on the market structure.[25] Because the Act became effective as of June 1, 1997, we use data for the period between July 1, 1997 and June 30, 2000. We obtain our main data on the branching of commercial banks from the Institutional Directory of the Federal Deposit Insurance Cooperation (FDIC). We augment the financial data of the banks with data from the Central Data Repository of the Federal Financial Institutions Examination Council. The data we construct contains information on the location of all branches and the financial statistics of every FDIC insured bank that had at least one branch in one of 13 states during the data period. The data also

[25] The 13 states are Arkansas, Colorado, Georgia, Iowa, Kentucky, Minnesota, Montana, Nebraska, New Jersey, North Dakota, Oklahoma, and Wyoming.

Market (# of obs = 723)	Mean	Std. Dev.	Min	Max
Population (thousand)	16.1	11.1	0.4	48.3
Income (per capita)	18,536.7	3,362.5	7.340	31,890
Total banks per market	38.2	17.4	4	91
Incumbents per market	4.9	3.0	1	25
Potential entrants per market	33.3	15.6	2	79
Operating banks per market	4.7	2.8	1	13
Number of mergers per market	0.6	0.8	0	5

Table 1: Descriptive Statistics — Market Level

keeps track of the banks' mergers and acquisitions. Data on merger contracts are obtained from the data set of SNL Financial, which reports deal values for each merger.[26]

Markets in the banking industry are known to be local in nature.[27] Existing works as well as antitrust analysis use geographic area as the definition of a market for the banking industry. Following Cohen and Mazzeo (2007) we focus our attention on rural markets, and we use a county as a market. This is because the typical market definition for urban areas (such as the Metropolitan Statistical Area) is likely to include submarkets within it. For this reason we exclude counties with a population greater than 50,000 from our data.[28] In such markets, consumers are also very less likely to use banks in other markets.

As discussed in the model section, we classify banks into incumbents and potential entrants in each market. We define banks that have operated in the market as of July 1, 1997 as incumbents. Regarding potential entrants, banks that have operated in a contiguous market during the data period are defined as potential entrants. There is a small number of banks that have entered though they are not identified as potential entrants according to this definition. Hence, we added these banks to the set of potential entrants as well, which also includes a small number of newly established banks that account for 0.1% of the potential entrants. We identify firm entry if the firm exists at the end of our sample period.

Table 1 reports the summary statistics of the market-level information. On average, there are 4.9 incumbent banks and 33.3 potential entrants in a market. There is substantial variation across markets for the number of incumbents and potential entrants. Among those firms, the number of operating banks is 4.7 on average. The average number of mergers per

[26] Merger data of SNL Financial do not necessarily have the same bank names as in FDIC data for the buyer and target banks. About a third of the mergers are matched by the FDIC certification number for both buyer and target banks. Another one third of the mergers are matched by the FDIC certification number of one side, and the holding company name and the FDIC holding company number. For the rest of the mergers, we matched manually using bank and holding company information from both data sets, and other sources such as regulatory filings. If a merger was still unmatched, we interpolated the transfer using buyer and target characteristics.

[27] See, e.g., Ruffer and Holcomb (2001), Ishii (2005), and Cohen and Mazzeo (2007).

[28] The number of markets does not change much if we use the criteria with a population less than 100,000.

	Mean	Std. Dev.	Min	Max
Incumbents				
Asset ($1M)	198.2	1,139.0	1.05	20,100
Deposit ($1M)	125.5	779.5	0	10,400
Equity Ratio	0.10	0.04	0.04	0.998
Potential Entrants				
Asset ($1M)	642.3	3,540.0	1.05	57,200
Deposit ($1M)	406.9	2,147.0	0	36,500
Equity Ratio	0.10	0.05	0.04	0.998

Table 2: Descriptive Statistics — Bank Characteristics.

market is 0.7.[29]

One of the assumptions of our empirical analysis is that the entry and merger decisions are independent across markets. Regarding this point, more than 80% of the incumbents were present only in one market, and more than 95% of the incumbents were present in less than three markets. Regarding actual entry, both the incumbents and potential entrants enter only one market in about 80% of the cases and less than three markets in more than 95% of the cases for incumbents and 92% for potential entrants. Conditioning on entry with merger, both types of banks enter less than three markets in 92% of the cases. Thus, in the vast majority of our data, banks do not overlap across markets, and we treat markets independently. The fact that our data is mostly from small regional banks in small regional markets helps us on the independence assumption.

Table 2 reports the summary statistics of the bank-level information. The incumbents' mean size of assets is much smaller than that of potential entrants at $198.2 million. This may reflect the fact that we define potential entrants as banks in contiguous markets, which tend to be larger than the market we consider. Descriptive statistics on the equity ratio are roughly the same for incumbents and potential entrants with a mean of 10%.

Table 3 provides the descriptive statistics of the same variables included in Table 2 for incumbent and potential-entrant banks that enter the market with a merger, respectively. Table 3 also reports the merger payment from buyer to target banks, which includes not only cash payment but also payment by share. As we saw in Table 2, the mean size of assets and deposits for incumbents is much smaller than that of potential entrants at $2,630 million and $1,770 million, respectively, while the equity ratio is almost the same. Comparing Tables 2 and 3, the average size of the assets and deposits are much larger for banks that

[29] During our sample period in our sample markets, there were no bank failure cases with open bank assistance by FDIC, while there were five bank failures where purchase and assumption (P&A) transactions were made. Since we observe such a small number of P&A transactions and cannot identify parameters specific to P&A, we treated these cases in the same way as other regular mergers.

	Mean	Std. Dev.	Min	Max
Incumbents				
Asset ($1M)	725.4	2,800.0	2.15	20,100
Deposit ($1M)	452.6	1570.5	0	10,400
Equity Ratio	0.09	0.05	0.04	0.998
Potential Entrants				
Asset ($1M)	2,630.0	8,620.0	2.15	57,200
Deposit ($1M)	1,770.0	5,540.0	0	36,500
Equity Ratio	0.09	0.05	0.04	0.998
Merger Payment ($1M)	517.1	2,435.6	0.6	21,237

Table 3: Descriptive Statistics — Bank Characteristics of the Merged Banks

enter the market with a merger, and also the standard deviations are much larger for those banks.

4 Identification and Estimation

In this section, we propose an estimation strategy for two-sided matching models based on moment inequalities and equalities. One of the major issues of estimating two-sided matching models is addressing the multiplicity of stable matchings. Our estimation strategy uses moment inequalities to deal with the issue of multiple equilibria similar to recent studies estimating noncooperative games by a set estimator (Ciliberto and Tamer, 2009, Ho, 2009, Kawai and Watanabe, forthcoming). We do so by exploiting the lattice structure of the set of equilibria (or $\mathcal{N}$-stable outcome, to be more precise). Before presenting our identification and estimation strategies, let us first describe the specification of the payoff function.

4.1 Specification

First, we provide a specification regarding the firm payoffs. In our environment profits depend on the outcome of the matching game as well as the firm and the market characteristics. As discussed in the model section, we follow Bresnahan and Reiss (1991) and assume that the number of operating firms, N, negatively affects the profit of the firm (note that N depends on entry and merger decisions jointly). The specific form of the profit function we consider for incumbent i (for entry without merger and no entry) is

$$\begin{aligned} \pi_i(i, \mathcal{N}_i) &= \alpha \mathcal{N}_i + z\beta_0 + x_i\beta_{1I} + \beta_{2I} + \xi_m, \\ \pi_i(o, \mathcal{N}_i) &= 0, \end{aligned}$$

where α is the degree of the negative externalities due to competition, and $\pi_i(k, \mathcal{N}_i) \equiv \pi_i(k, \mathcal{N}_i(k))$ with slight abuse of notation. Market characteristics, which include population and average income, are denoted by z and the effect of market characteristics on profit is captured by β_0. The effect of firm i's characteristics, x_i, in the case of entry without merger is denoted by β_1. Regarding firm-specific characteristics, we consider each bank's equity ratio and asset size. The next term, β_{2I}, is a constant term for incumbents entering without merger. The last term, ξ_m, denotes market-level profit shock, which is an i.i.d. draw from a normal distribution, $N(0, 1)$. This is for a normalization. If the firm does not enter the market, the profit is zero. Similarly, we write the profit function for entrant e (for entry without merger and no entry) as

$$\begin{aligned} \pi_e(e, \mathcal{N}_e) &= \alpha \mathcal{N}_e + z\beta_0 + x_e\beta_{1E} + \beta_{2E} + \xi_m, \\ \pi_e(o, \mathcal{N}_e) &= 0, \end{aligned}$$

where x_e denotes firm e's characteristics, and β_{2E} is a constant term for potential entrants, which we allow to differ from β_{2I} for incumbents. As in the case of incumbents, firm e's characteristics (x_e) include the equity ratio and the asset size. The difference between β_{2E} and β_{2I} captures the difference of entry costs across potential entrants and incumbents, if they enter the market without merger.

Next, in order to write the profit for the case of entry with merger, we start with the profit of the merged entity. The profit after merger between firms e and i given the (dis)synergy function f and the estimation by firm $j \in \{e, i\}$, $\pi(k_{ei}, \mathcal{N}_j)$, is written as

$$\begin{aligned} \pi(k_{ei}, \mathcal{N}_j) &= \alpha \mathcal{N}_j + z\beta_0 + f(x_i, x_e, x_{ie}) + \xi_m, \\ f(x_i, x_e, x_{ie}) &= \beta_{2M} + x_i\beta_3 + x_e\beta_4 + x_i x_e \beta_5 + x_{ie}\beta_6, \end{aligned}$$

where the synergy or dissynergy of a match is represented by a synergy function $f(x_i, x_e, x_{ie})$, which depends on firm i's characteristics (x_i), firm e's characteristics (x_e), and match-specific characteristics (x_{ie}). For match specific characteristics, we use the distance between the headquarters of two banks and the indicator variable for the same bank holding company. β_{2M} is a constant term for (dis)synergy, β_3 and β_4 are the effects of the incumbent's and potential entrant's characteristics on (dis)synergy of the merger, respectively, and β_5 captures the effect of the interaction terms of both firms' characteristics on (dis)synergy. β_6 measures how the match-specific characteristics such as the distance between headquarters affect the (dis)synergies.[30] Note that we do not consider (dis)synergies across markets

[30] Note that the match specific characteristics, x_{ie}, affect the synergy function, but not the payoff when the two firms do not merge. We use this exclusion restriction to identify the synergy function, on top of the exclusion restriction of x_j, both of which we discuss in the next section.

although such factors may be present in some industries. This is because vast majority of banks are present only in one market in our data as described in the data section.

Because the terms of merger contracts take cash and stock as medium of payment, we consider the space of the term of trade P as $P = T \times R$, where $T = \{\underline{t}, ..., 0, ..., \bar{t}\}$ corresponds to a finite set of cash transfers, and $R = \{0, ..., 1\}$ corresponds to a finite set of stock shares between e and i after merger. Now we can write the profit function for the case of mergers with contract k_{ei} for firms e and i as

$$\begin{aligned} \pi_e(k_{ei}, \mathcal{N}_e) &= r_{ei}\pi(k_{ei}, \mathcal{N}_e) - t_{ei}, \\ \pi_i(k_{ei}, \mathcal{N}_i) &= (1 - r_{ei})\pi(k_{ei}, \mathcal{N}_i) + t_{ei}, \end{aligned}$$

where $t_{ei} \in T$ denotes a cash transfer from e to i and $r_{ei} \in R$ denotes a payment in stock shares of the merged entity from e to i. Note that both r_{ei} and t_{ei} are observable for the realized merger contracts as data.

Finally, we specify the payoff of each firm considering both profit and other factors. As discussed in the model section, we allow idiosyncratic factors to affect entry and merger decisions because these factors such as the degree of shareholder control, CEOs' tastes and abilities may play an important role as illustrated in Jensen and Meckling (1976), Malmendier and Tate (2008) and Goldfarb and Xiao (2011).

We consider such idiosyncratic shock ε_{ie} for firm i merging with firm e, and ε_{ii} for firm i entering without merger. These idiosyncratic shocks are unobservable to the econometrician though they are observable to the firms. Similarly we have ε_{ei} for firm e merging with firm ι, and ε_{ee} for firm e entering without merger. These shocks are i.i.d. draws from the standard normal distribution.[31] Furthermore, the relative importance of shocks ε_{ie} and ε_{ei} to the monetary profit are allowed to be different across banks. Specifically, we capture the relative importance of the idiosyncratic shocks to π_j by σ_I for incumbents and by σ_E for potential entrants. We thus write the payoff of incumbent i as

$$\begin{aligned} \Pi_i(k_{ei}, \mathcal{N}_i) &= \frac{1}{\sigma_I}\pi_i(k_{ei}, \mathcal{N}_i) + \varepsilon_{ie}, \\ \Pi_i(i, \mathcal{N}_i) &= \frac{1}{\sigma_I}\pi_i(i, \mathcal{N}_i) + \varepsilon_{ii}, \\ \Pi_i(o, \mathcal{N}_i) &= 0, \end{aligned}$$

[31] Additive separability and i.i.d. normality assumptions are not crucial for identification and estimation, and we make these for computational simplicity.

and the payoff of potential entrant e as

$$\begin{aligned}
\Pi_e(k_{ei}, \mathcal{N}_e) &= \frac{1}{\sigma_E}\pi_e(k_{ei}, \mathcal{N}_e) + \varepsilon_{ei}, \\
\Pi_e(e, \mathcal{N}_e) &= \frac{1}{\sigma_E}\pi_e(e, \mathcal{N}_e) + \varepsilon_{ee}, \\
\Pi_e(o, \mathcal{N}_e) &= 0.
\end{aligned}$$

For notational convenience, we define $\theta \equiv (\alpha, \beta_0', \beta_1', \beta_{2I}, \beta_{2E}, \beta_{2M}, \beta_3', \beta_4', \beta_5', \beta_6', \sigma_I, \sigma_E)$, and the space of θ as Θ.

4.2 Identification

4.2.1 Identified Set

Our identification is based on the restrictions provided by the equalities (1) and (2) in Corollary 1 and the inequalities (3) and (4) in Corollary 2. In Corollary 2 we show that the payoffs in the two extremum equilibria (or $\mathcal{N}$-stable outcomes to be more precise), $\overline{K}^{*\mathcal{E}}$ and $\overline{K}^{\mathcal{I}*}$, are the upper and lower bounds of any equilibrium payoff. Hence they constitute the bounds of the payoffs corresponding to the observed data. Furthermore, Corollary 1 shows that, in any $\mathcal{N}$-stable outcome, the number of operating firms is the same. Also, the payoffs of firms that do not merge are the same in any $\mathcal{N}$-stable outcome. These results lead us to construct moment equalities regarding the number of operating firms and the payoffs of non-merging firms. Hence the observed data should match the number of operating firms and the payoffs of non-merging firms in $\overline{K}^{*\mathcal{E}}$ and $\overline{K}^{\mathcal{I}*}$.

Let us first revisit the result of Corollary 2 (more specifically, equations (3) and (4)): the two extremum $\mathcal{N}$-stable outcomes provide the highest payoff to one side of the banks and the lowest payoff to the other side, i.e., (suppressing the dependence on $\mathcal{N}$ for notational convenience) for any $\mathcal{N}$-stable outcome $\overline{K}^*$,

$$\begin{aligned}
\Pi_e(\overline{K}^{*\mathcal{E}}; \theta) &\geq \Pi_e(\overline{K}^*; \theta) \geq \Pi_e(\overline{K}^{*\mathcal{I}}; \theta), \ \forall e \in \mathcal{E}, \\
\Pi_i(\overline{K}^{*\mathcal{I}}; \theta) &\geq \Pi_i(\overline{K}^*; \theta) \geq \Pi_i(\overline{K}^{*\mathcal{E}}; \theta), \ \forall i \in \mathcal{I}.
\end{aligned}$$

In other words, given θ, e's payoff in any $\mathcal{N}$-stable outcome is bounded above by the payoff in the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome, $\Pi_e(\overline{K}^{*\mathcal{E}}; \theta)$, and bounded below by the payoff in the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome, $\Pi_i(\overline{K}^{*\mathcal{I}}; \theta)$. Similarly the payoffs for all incumbents are bounded above and below by $\Pi_i(\overline{K}^{*\mathcal{I}}; \theta)$ and $\Pi_i(\overline{K}^{*\mathcal{E}}; \theta)$, respectively. Observe that we can compute these bounds given θ (and shocks) using the algorithm shown in Section 2.3.

We can also compute the payoffs for each firm j corresponding to the data, $\Pi_j(\overline{K}^{*DATA}; \theta)$, given θ (and shocks), where we denote the particular equilibrium selected in the observation

as $\overline{K}^{*DATA}$. Though we cannot know which equilibrium the data-generating process corresponds to, the payoff corresponding to the equilibrium in the data must still be bounded by $\Pi_j(\overline{K}^{*\mathcal{I}};\theta)$ and $\Pi_j(\overline{K}^{*\mathcal{E}};\theta)$ for all j. Hence, we can consider the following types of inequalities:

$$E\left[\Pi_e(\overline{K}^{*\mathcal{E}};\theta) - \Pi_e(\overline{K}^{*DATA};\theta)|X\right] \geq 0, \tag{5}$$

$$E\left[\Pi_i(\overline{K}^{*\mathcal{I}};\theta) - \Pi_i(\overline{K}^{*DATA};\theta)|X\right] \geq 0, \tag{6}$$

$$E\left[\Pi_e(\overline{K}^{*DATA};\theta) - \Pi_e(\overline{K}^{*\mathcal{I}};\theta)|X\right] \geq 0, \tag{7}$$

$$E\left[\Pi_i(\overline{K}^{*DATA};\theta) - \Pi_i(\overline{K}^{*\mathcal{E}};\theta)|X\right] \geq 0, \tag{8}$$

where $X = \{(x_{i_1}, ..., x_{i_{N_\mathcal{I}}}), (x_{e_1}, ..., x_{e_{N_\mathcal{E}}}), z\}$ denotes firm characteristics of all firms in a market and market characteristics. Note that these inequalities are at the level of individual firms. However, our unit of observation is a market, and identity and number of incumbents and potential entrants differ across markets. Thus, we use moments based on these inequalities at market level in our estimatio, which we describe in Appendix B (we construct 44 moment inequalities.).

Next, we discuss moment equalities resulting from Corollary 1 (more specifically, equations (1) and (2)). Part (iii) of Corollary 1 shows that the number of operating firms and the number of mergers are identical in any $\mathcal{N}$-stable outcome (equation (2)). Though the econometrician cannot know which equilibrium the data-generating process corresponds to, the number of operating firms and the number of mergers are the same in any equilibrium (or in any $\mathcal{N}$-stable outcome). Thus, the observed data on the numbers of operating firms and mergers equal those in the two extremum $\mathcal{N}$-stable outcomes, $\overline{K}^{*\mathcal{E}}$ and $\overline{K}^{\mathcal{I}*}$, and we can write the following equalities:

$$\begin{aligned} E\left[N^{DATA}_{merge}|X\right] &= E\left[\gamma(\overline{K}^{*\mathcal{E}}, N^{DATA};\theta)|X\right] = E\left[\gamma(\overline{K}^{*\mathcal{I}}, N^{DATA};\theta)|X\right], \\ E\left[N^{DATA}|X\right] &= E\left[\lambda(\overline{K}^{*\mathcal{E}}, N^{DATA};\theta)|X\right] = E\left[\lambda(\overline{K}^{*\mathcal{I}}, N^{DATA};\theta)|X\right], \end{aligned} \tag{9}$$

where N^{DATA}_{merge} and N^{DATA} are the numbers of mergers and operating firms in the observed data. Regarding Part (ii) of Corollary 1, we have equation (1) which states that unmatched (non-merging) firms earn exactly the same payoff in any $\mathcal{N}$-stable outcome, i.e., for any j choosing either $\{j\}$ or $\{o\}$, we have

$$E\left[\Pi_j(\overline{K}^{*DATA};\theta)|X\right] = E\left[\Pi_j(\overline{K}^{*\mathcal{E}};\theta)|X\right] = E\left[\Pi_j(\overline{K}^{*\mathcal{I}};\theta)|X\right]. \tag{10}$$

Same as moment inequalities, these equalities are at the individual firm level, and we de-

scribe the moments we use in our estimation based on these equalities in Appendix B (we construct 14 moment equalities.).

Finally, we define the identified set Θ_{id} using both moment inequalities and equalities as $\Theta_{id} = \{\theta \in \Theta :$ inequalities (5)–(8) and equalities (9) and (10) are satisfied at $\theta\}$.

4.2.2 Exclusion Restrictions

Our identification depends on two types of exclusion restrictions: the exclusion restriction employed in the regular entry model and the exclusion restriction for the synergy function.[32] We need these two types of exclusion restrictions because x_i and x_e affect not only the payoff of entering without merger but also the payoff of entering with merger.

The first type of exclusion restriction we use is the one adopted in the literature on estimating entry models. As in Berry (1992) and Tamer (2003), we need a variable that affects a firm's profit but does not enter the other firms' profit functions. In our model, the asset size of a bank, x_j^{asset}, serves as an exclusion restriction of this type: In the banking literature, demand is typically modeled as a function of geographic proximity to the banking facility and interest rate (see, e.g., Ishii, 2005), while the asset size is an important factor on the cost side (see, e.g., McAllister and McManus, 1993).

The second type of exclusion restriction we use concerns the identification of the synergy function. To identify synergy function f, the first type of exclusion restriction does not suffice if all exogenous variables in f are also included in $\Pi_j(j, N)$ (payoff for entry without merger). Thus, we require a match-specific variable that enters the synergy function, but affects neither the (dis)synergy of any other combination of firms nor the payoffs of the two firms entering without merger.

In our specification, we use the distance between the headquarters of the incumbent and the potential entrant, which is the first element of x_{ie}, $x_{ie}^{(1)}$. This variable affects the post-merger synergy for various reasons, such as communication between the workers of the target and acquiring banks, while it is unlikely to affect the payoff of entry without merger and mergers of any other bank pairs.

Our identification argument proceeds in two steps using the two types of exclusion restrictions in each step. First, we use the second type of exclusion restriction to identify all the parameters that are not included in the synergy function. The assumption we need is that $f(x_i, x_e, x_{ie}) \to -\infty$ as $x_{ie}^{(1)} \to \infty$, i.e., the dissynergy goes to infinity as the distance becomes infinity. This implies that as $x_{ie}^{(1)} \to \infty$ we obtain $\Pi_e(k_{ei}, N) \to -\infty$ and

[32] Our identification uses identification at infinity arguments that are commonly used. See, e.g., Ciliberto and Tamer (2009) for similar identification arguments for entry models. Though our estimation strategy is robust to lack of point-identification, it is useful to discuss how the model can be point-identified at infinity. Our argument shows that the model can be point-identified under certain conditions and the identified set becomes sharper to the extent such conditions are satisfied in the data.

$\Pi_i(k_{ei}, N) \to -\infty$ without changing $\Pi_j(j, N)$, i.e., the payoff with merger becomes strictly less than that of entering without merger.

Now, given that the firms have no incentive to choose mergers ($\Pi_e(k_{ei}, N) < \Pi_e(o, N) = 0$ and $\Pi_i(k_{ei}, N) < \Pi_i(o, N) = 0$), the game is equivalent to a regular entry model. This is because we can ignore entry with merger in such a case (it gives strictly lower payoff than not entering). Therefore, $\Pi_j(j, N)$ is identified by the first type of exclusion restriction as shown in Berry (1992) and Tamer (2003).

Second, we discuss the identification of the synergy function. Given that $\Pi_j(j, N)$ is identified in the first step, we can use the variation of outcome k_{ei} and that of characteristics x_i, x_e, and x_{ie} to identify $\Pi_i(k_{ei}, N)$ and $\Pi_e(k_{ei}, N)$. Because the effect of x_i and x_e on $\Pi_j(j, N)$ is identified in the first step, we can isolate the effects of x_i and x_e on $\Pi_i(k_{ie}, N)$ from those on $\Pi_i(i, N)$. Same argument applies to $\Pi_e(k_{ie}, N)$ and $\Pi_e(e, N)$. Finally, as we have isolated the effects of x_i and x_e on $\Pi_i(k_{ie}, N)$ and $\Pi_e(k_{ie}, N)$, the variation of k_{ie} and that of x_i, x_e, and x_{ie} identify the (dis)synergy function f.

4.3 Estimation

Following the identification argument, we estimate the model using the moment inequality estimator developed by Andrews and Soares (2010). If an econometrician knew the equilibrium selection mechanism, a single outcome would correspond to one realization of the unobserved error terms (ξ, ε). In such a case, we could employ estimation procedures such as GMM or MLE. However, as discussed in Section 4.2, the multiplicity of equilibria ($\mathcal{N}$-stable outcomes in our case) implies that the model parameters are only partially identified: This makes the use of a set estimator more appropriate.

We denote all 44 moment inequalities and 14 equalities discussed in Section 4.2 (and described in Appendix B) by

$$\begin{aligned} E\left[h_l(X;\theta)\right] &\geq 0, \quad l = 1, ..., 44 \\ E\left[h_l(X;\theta)\right] &= 0, \quad l = 45, ...58 \end{aligned}$$

Now we describe our estimation procedure using these moments. Our procedure solves the generalized Gale-Shapley algorithm with externalities for each simulation draw, and computes the sample analogue of moment inequalities and equalities in the following way.

1. Fix parameter θ. For each market $m = 1, ..., M$, obtain S draws of $\boldsymbol{\varepsilon}_{\mathcal{I}}^{ms} = \{\{\varepsilon_{ie}^{ms}\}_{i=1}^{N_i^m}\}_{e=1}^{N_e^m}$, $\boldsymbol{\varepsilon}_{\mathcal{E}}^{ms} = \{\{\varepsilon_{ei}^{ms}\}_{e=1}^{N_e^m}\}_{i=1}^{N_i^m}$, and ξ_m^s from distributions $g_{\mathcal{I}}$, $g_{\mathcal{E}}$, and g_m.

2. For each draw $\boldsymbol{\eta}^{ms} = (\boldsymbol{\varepsilon}_{\mathcal{I}}^{ms}, \boldsymbol{\varepsilon}_{\mathcal{E}}^{ms}, \xi_m^s)$ in each market m, run an $\mathcal{E}$-proposing generalized Gale-Shapley algorithm with externalities to obtain the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome.

Run also an $\mathcal{I}$-proposing generalized Gale-Shapley algorithm with externalities for the same draw $(\boldsymbol{\varepsilon}_{\mathcal{I}}^{ms}, \boldsymbol{\varepsilon}_{\mathcal{E}}^{ms}, \xi_m^s)$ in the same market m to obtain the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcomes.

3. Construct a sample analogue of the moment inequalities and equalities using the $\mathcal{E}$-optimal and $\mathcal{I}$-optimal stable outcomes as well as the observed match $\overline{K}^{*DATA}$:

$$\begin{aligned} \frac{1}{MS}\sum_{m=1}^{M}\sum_{s=1}^{S} h_l(X^m, \boldsymbol{\eta}^{ms}; \theta) &\geq 0, \quad l = 1, ..., 44 \\ \frac{1}{MS}\sum_{m=1}^{M}\sum_{s=1}^{S} h_l(X^m, \boldsymbol{\eta}^{ms}; \theta) &= 0, \quad l = 45, ..., 58. \end{aligned}$$

4. Use the moment inequality estimator by Andrews and Soares (2010).

The specific functions we use to construct test statistics and a critical value in Andrews and Soares (2010) are $S = S_1$ and $\varphi_j = \varphi_j^{(4)}$ with the number of bootstrapping $R = 1000$. Because we cannot report a 21-dimensional confidence set, we compute min and max of the confidence set projected on each dimension and report it in the next section (see Appendix C for details).

5 Results and Counterfactual Experiments

5.1 Parameter Estimates

The confidence intervals for the parameters are reported in Table 4. The exact specification of the payoff function for incumbent i is

$$\begin{aligned} \Pi_i(o, N) &= 0, \\ \Pi_i(i, N) &= \frac{\alpha N + \mathbf{z}[\beta_0^{pop}, \beta_0^{income}] + \mathbf{x}_i[\beta_{1I}^{size}, \beta_{1I}^{eq\_ratio}] + \beta_{2I} + \xi_m}{\sigma_I} + \varepsilon_{ii}, \\ \Pi_i(k_{ei}, N) &= \frac{(1 - r_{ei})\left(\alpha N + \mathbf{z}[\beta_0^{pop}, \beta_0^{income}] + f(\mathbf{x}_i, \mathbf{x}_e, \mathbf{x}_{ie}) + \xi_m\right) + t_{ei}}{\sigma_I} + \varepsilon_{ie}, \end{aligned}$$

where $\xi_m \sim N(0,1)$, and for potential entrant e is

$$\begin{aligned}
\Pi_e(o,N) &= 0, \\
\Pi_e(e,N) &= \frac{\alpha N + \mathbf{z}[\beta_0^{pop}, \beta_0^{income}] + \mathbf{x}_e[\beta_{1E}^{size}, \beta_{1E}^{eq_ratio}] + \beta_{2E} + \xi_m}{\sigma_E} + \varepsilon_{ee}, \\
\Pi_e(k_{ei},N) &= \frac{r_{ei}\left(\alpha N + \mathbf{z}[\beta_0^{pop}, \beta_0^{income}] + f(\mathbf{x}_i, \mathbf{x}_e, \mathbf{x}_{ie}) + \xi_m\right) - t_{ei}}{\sigma_E} + \varepsilon_{ei},
\end{aligned}$$

where the synergy function in the payoff is specified as

$$\begin{aligned}
f(\mathbf{x}_i, \mathbf{x}_e, \mathbf{x}_{ie}) &= \beta_{2M} + \mathbf{x}_i[\beta_3^{size}, \beta_3^{eq_ratio}] + \mathbf{x}_e[\beta_4^{size}, \beta_4^{eq_ratio}] \\
&\quad + \mathbf{x}_i^1 \mathbf{x}_e^1 \beta_5^{size} + \mathbf{x}_i^2 \mathbf{x}_e^2 \beta_5^{eq_ratio} + d_{ie}\beta_{6,dist} + d_{ie}^2 \beta_{6,dist2} + h_{ie}\beta_{6,bhc},
\end{aligned}$$

where $\mathbf{x}_{ie}$ consists of the distance between the headquarter of i and e, d_{ie}, and the indicator variable for the same bank holding company, h_{ie}, i.e. $\mathbf{x}_{ie} = [d_{ie}, h_{ie}]'$. $\mathbf{z}$ is a vector of the market characteristics (log of population and log of per capita income), $\mathbf{x}_i$ is incumbent i's characteristics (log of total asset size and equity ratio), $\mathbf{x}_e$ is potential entrant e's characteristics (log of total asset size and equity ratio), $\mathbf{x}_i^1 \mathbf{x}_e^1$ is an interaction term of incumbent i and potential entrant j's asset size, $\mathbf{x}_i^2 \mathbf{x}_e^2$ is an interaction term of incumbent i and potential entrant j's equity ratio, and r_{ei} and t_{ei} are the terms of the merger contract k_{ei} (payment by stock and cash). Note that β_{2I} and β_{2E} capture the size of entry barriers of incumbents and potential entrants for entry without merger, respectively, and β_{2M} measures the cost of entry by merger.

First, we discuss the estimates for the parameters in (dis-)synergy function f. The estimates for the effect of asset size for incumbents and potential entrants are $\beta_3^{size} = [-0.048, -0.018]$ and $\beta_4^{size} = [0.332, 0.406]$. That is, smaller incumbents and larger potential entrants have higher synergies. The interaction term for the incumbents' and potential entrants' asset size is $\beta_5^{size} = [-0.028, -0.023]$, which also implies that the smaller the incumbent and the larger the potential entrant, the higher is the synergy. The effects of the distance between two merging banks are $\beta_{6,dist2} = [-0.027, -0.006]$ and $\beta_{6,dist} = [0.792, 3.187]$ which indicates that the synergy becomes smaller as the distance between the merging firms becomes greater. The estimates confirm the pattern that small regional banks are acquired by relatively larger regional banks as our observation is focused on small regional markets.

The estimates for the synergy effect of the equity ratio exhibit a similar pattern. The results show that the equity ratio affects the synergy differently for incumbents and potential entrants based on the estimates of $\beta_3^{eq\_ratio} = [-3.526, -1.085]$ and $\beta_4^{eq\_ratio} =$

Parameter	Confidence Interval	Parameter	Confidence Interval
α	$[-1.763, -1.443]$	$\beta_3^{eq\_ratio}$	$[-3.526, -1.085]$
β_0^{pop}	$[1.007, 1.231]$	β_3^{size}	$[-0.048, -0.018]$
β_0^{income}	$[0.227, 0.277]$	$\beta_4^{eq\_ratio}$	$[2.326, 5.235]$
$\beta_{1I}^{eq\_ratio}$	$[0.115, 1.908]$	β_4^{size}	$[0.332, 0.406]$
β_{1I}^{size}	$[-0.201, -0.109]$	$\beta_5^{eq\_ratio}$	$[-22.598, -10.272]$
$\beta_{1E}^{eq\_ratio}$	$[0.367, 0.882]$	β_5^{size}	$[-0.028, -0.023]$
β_{1E}^{size}	$[-0.813, -0.665]$	$\beta_{6,dist2}$	$[-0.027, -0.006]$
β_{2I}	$[2.853, 5.705]$	$\beta_{6,dist}$	$[0.792, 3.187]$
β_{2E}	$[-8.651, -4.866]$	β_{bhc}	$[2.452, 7.765]$
β_{2M}	$[-1.912, -1.434]$	σ_I	$[0.632, 0.666]$
		σ_E	$[0.747, 0.919]$

Table 4: Confidence Intervals. We report the 95% confidence intervals calculated following Andrews and Soares (2010).

$[2.326, 5.235]$. Incumbents with lower equity ratio have higher synergy, while a high equity ratio of a potential entrant impacts synergy positively. The estimate for the interaction term of equity ratio, $\beta_5^{eq\_ratio} = [-22.598, -10.272]$, implies that synergy is further strengthened if the equity ratio of the incumbent is low and that of the potential entrant is high. This result implies that less financially sound incumbents to are more likely to have higher synergies with healthy potential entrant banks.

Constant terms for incumbents and potential entrants are estimated as $\beta_{2E} = [-8.651, -4.866]$ and $\beta_{2I} = [2.853, 5.705]$, while that for the entry by merger is $\beta_{2M} = [-1.912, -1.434]$. These constant terms are interpreted as sunk costs of entry for different forms of entry (entry with or without merger) for potential entrants and incumbents. Potential entrants have much higher entry costs if they enter de novo instead of entry by merger ($\beta_{2E} < \beta_{2M}$), which partly explains why potential entrants tend to enter by merger rather than de novo. Also, incumbents have much lower entry costs compared with potential entrants. These effects are significantly large, and are comparable to the competition effect that is estimated as $\alpha = [-1.763, -1.443]$.

Regarding bank characteristics, the estimates for the effect of equity ratio and asset size are $\beta_{1I}^{eq\_ratio} = [0.115, 1.908]$ and $\beta_{1E}^{eq\_ratio} = [0.367, 0.882]$, and $\beta_{1I}^{size} = [-0.201, -0.109]$ and $\beta_{1E}^{size} = [-0.813, -0.665]$, implying that a healthy balance sheet has strong effects on profitability and that a larger bank tend to be less profitable in small regional markets our data is from. Concerning market characteristics, the coefficients for both per capita income and population are positive. This implies that banks' profits tend to be higher in markets with a larger population and higher income ceteris paribus.

	Mean	Std Dev	10%-tile	Median	90%-tile
Number of Operating Firms					
Data	4.668	2.812	2	4	9
Counterfactual	$[4.039, 4.044]$	$[2.143, 2.144]$	$[1, 1]$	$[4, 4]$	$[7, 7]$
Number of Entry by Merger					
Data	0.598	0.815	0	0	2
Counterfactual	$[0.657, 0.757]$	$[0.965, 1.030]$	$[0, 0]$	$[0, 1]$	$[2, 2]$

Table 5: Counterfactual Experiment: Effects of the Riegle-Neal Act. The counterfactual experiment prohibits entry without merger for potential entrants.

Lastly, we discuss the estimates of the relative importance of idiosyncratic shocks on the payoffs. The estimates of the standard deviation are $\sigma_I = [0.632, 0.666]$ and $\sigma_E = [0.747, 0.919]$ compared to the market-level shock that is normalized to have standard deviation of one. This implies that idiosyncratic shocks are larger for potential entrants.

5.2 Counterfactual Experiment

We study the effects of the Riegle-Neal Act by conducting a counterfactual experiment. The Riegle-Neal Act deregulated the intrastate de novo entry for the 13 states that had not deregulated this form of entry at the time of the Act, in addition to deregulating interstate banking. Our data correspond to the period that is right after the deregulation of de novo entry by the Act. In our counterfactual, we simulate market structure by prohibiting de novo entry for potential entrants. The difference between the data and the predicted outcome of the counterfactual experiment presents the effect of the Act on the market structure.

Table 5 reports the results of the counterfactual experiment. We obtain our results by simulating the model without the choice of de novo entry for potential entrants (setting $\beta_{2E} = -\infty$) and letting the parameter values to move within the confidence set. Computational details are discussed in Appendix D. The results show that the number of operating banks would have been smaller if de novo entry was prohibited. As the last two rows of Table 5 show, the mean number of entry by merger is increased under the counterfactual. This is because prohibiting de novo entry by potential entrants increases the incentive for the potential entrants to enter by merger. However, the increase of entry by merger is not large enough to undo the decrease of de novo entry, resulting in smaller number of operating banks on average.

6 Conclusion

We study entry and merger decisions of banks jointly in this paper. We show the existence of the stable outcome in a two-sided matching model with externalities by proposing an algorithm. Using data on commercial banks in the U.S., we then estimate the model with a moment inequalities estimator based on the equilibrium characterization of the matching model without imposing an equilibrium selection mechanism. We find that entry barriers differ significantly across modes of entry and that synergy is larger when incumbent banks have a less healthy balance sheet and smaller size and potential entrants have larger asset size and healthier balance sheet.

There are many issues left for a future research regarding firms' merger and entry decisions. One issue we could not address was how entry by merger affects industry dynamics. A natural step would be to consider a merger as an additional investment tool in a dynamic industry competition model. Another issue concerns the way the externalities of post-entry competition affect firms. One can extend the model to consider the market with vertical or horizontal differentiations as in Mazzeo (2002) and Seim (2006).

References

[1] Adachi, Hiroyuki (2000) "On a Characterization of Stable Matchings," *Economics Letters*, Vol. 68, pp.43–49.

[2] Akkus, Oktay and Alī Hortaçsu (2007) "The Determinants of Bank Mergers: A Revealed Preference Analysis," mimeo.

[3] Andrews, D. W. K. and Soares, G. (2010) "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, pp. 119–157

[4] Baccara, Mariagiovanna, Ayse Imrohoroglu, Alistair J. Wilson, and Leeat Yariv (2012) "A Field Study on Matching with Network Externalities," *American Economic Review*, Vol 102(5), pp. 1773-1804.

[5] Berry, Steven (1992) "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, Vol. 60 (4), pp. 889-917.

[6] Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff (forthcoming) "Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Disentangling the Preferences of Teachers and Employers," *Journal of Labor Economics*.

[7] Bresnahan, Timothy F. and Peter C. Reiss (1991) "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, Vol. 99, pp. 977-1009.

[8] Chiappori, Pierre-André, Bernerd Salanie and Yoram Weiss (2010), "Partner Choice and the Marital College Premium," mimeo.

[9] Ciliberto, Federico and Elie Tamer (2009) "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, Vol. 77 (6), pp. 1791-1828.

[10] Choo, Eugene and Aloysius Siow (2006) "Who Marries Whom and Why," *Journal of Political Economy*, Vol.114 (1), pp. 175–201.

[11] Cohen, Andrew and Michael Mazzeo (2007) "Market Structure and Competition among Retail Depository Institutions," *Review of Economics and Statistics*, Vol. 89, pp. 60-74.

[12] Echenique, Federico and Jorge Oviedo (2006) "A Theory of Stability in Many-to-Many Matching Markets," *Theoretical Economics*, Vol.1, pp. 233-273.

[13] Echenique, Federico, SangMok Lee, Matthew Shum, and M. Bumin Yenmez (forthcoming) "The Revealed Preference Theory of Stable and Extremal Stable Matchings," *Econometrica*.

[14] Fox, Jeremy T. (2010) "Identification in Matching Games," *Quantitative Economics*, Vol. 1, pp. 203-254.

[15] Galichon, Alfred, and Bernard Salanié (2010) "Matching with Trade-offs: Revealed Preferences over Competing Characteristics," mimeo.

[16] Goldfarb, Avi, and Mo Xiao (2011) "Who Thinks about the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets," *American Economic Review*, Vol. 101(7), pp. 3130-3161.

[17] Gordon, Nora and Brian Knight (2009) "A Spatial Merger Estimator with an Application to School District Consolidation," *Journal of Public Economics*, Vol. 93, pp. 752-765.

[18] Gowrisankaran, Gautam (1999) "A Dynamic Model of Endogenous Horizontal Mergers," *RAND Journal of Economics*, Vol. 30, pp. 56-83.

[19] Hafalir, Isa E. (2008) "Stability of Marriage with Externalities," *International Journal of Game Theory*, Vol. 37 (3), pp. 353-369.

[20] Hatfield, John William and Paul R. Milgrom (2005) "Matching with Contracts," *American Economic Review*, Vol. 95 (4), pp. 913-935.

[21] Hennart, Jean-Francois and Young-Ryeol Park (1993) "Greenfield vs. Acquisition: The Strategy of Japanese Investors in the United States," *Management Science*, Vol. 39(9), pp. 1054-1070.

[22] Hsieh, Yu-Wei (2011) "Understanding Mate Preferences from Two-Sided Matching Matkets: Identification, Estimation and Policy Analysis," mimeo.

[23] Ho, Kate (2009) "Insurer-Provider Networks in the Medical Care Market," *American Economic Review*, Vol. 99 (1), pp. 393-430.

[24] Ishii, Joy (2005) "Compatibility, Competition, and Investment in Network Industries: ATM Networkds in the Banking Industry," mimeo.

[25] Jensen, M. C. and W. H. Meckling (1976) "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, Vol, 3 (4), pp. 305–360.

[26] Jia, Panle (2008) "What Happens When Wal-mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, Vol. 76 (6), pp. 1263-1316.

[27] Kamien, Morton I., and Israel Zang (1990) "The Limits of Monopolization Through Acquisition," *Quarterly Journal of Economics*, Vol. 105 (2), pp. 465-499.

[28] Kawai, Kei, and Yasutora Watanabe, (forthcoming) "Inferring Strategic Voting," *American Economic Review.*

[29] Malmendier, Ulrike, and Geoffrey Tate (2008) "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics,* Vol. 89(1), pp. 20-43.

[30] Mazzeo, Michael J. (2002) "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, Vol 33(2), pp.1–22.

[31] McAllister, P. H. and Douglas McManus (1993) "Resolving the Scale Efficiency Puzzle in Banking," *Journal of Banking and Finance*, Vol. 17, pp.389-405.

[32] Nishida, Mitsukuni (2012) "Estimating a Model of Strategic Network Choice: the Convenience-Store Industry in Okinawa," mimeo.

[33] Nocke, Volker and Michael D. Whinston (2010) "Dynamic Merger Review," *Journal Political Economy*, Vol. 118, pp. 1200-51.

[34] Nocke, Volker and Stephen Yeaple (2007) "Cross-Border Mergers and Acquisitions vs. Greenfield Foreign Direct Investment: The Role of Firm Heterogeneiry," *Journal of International Economics*, Vol. 72, pp. 336-365.

[35] Ostrovsky, Michael (2008) "Stability in Supply Chain Networks," *American Economic Review*, Vol. 98 (3), pp. 897-923.

[36] Park, Minjung (2011) "An Empirical Two-Sided Matching Model of Acquisitions: Understanding Merger Incentives and Outcomes in the Mutual Fund Industry," mimeo.

[37] Perez-Saiz, Hector (2012) "Building New Plants or Entering by Acquisition? Firm Heterogeneity, Identification and Entry Barriers," mimeo.

[38] Pesendorfer, Martin (2005) "Mergers under Entry," *RAND Journal of Economics*, Vol. 36 (3), pp. 661-679.

[39] Qiu, Larry D., and Wen Zhou (2007) "Merger Waves: A Model of Endogenous Mergers," *RAND Journal of Economics*, Vol. 38 (1), pp. 214-226.

[40] Roth, Alvin E. (1986) "On the Allocation of Residents to Rural Hospitals: a General Property of Two-Sided Matching Markets," *Econometrica,* Vol. 54 (2), pp.425-427.

[41] Roth, Alvin E. (2008) "Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions," *International Journal of Game Theory,* Vol 36, pp. 537-569.

[42] Ruffer, Rochelle L. and John P. Holcomb Jr. (2001) "To Build or Buy: An Empirical Study of the Characteristics Affecting a Bank's Expansion Decisions," *Journal of Economics & Business*, Vol. 53, pp. 481-495.

[43] Sasaki, Hiroo and Manabu Toda (1996) "Two-Sided Matching Problems with Externalities," *Journal of Economic Theory*, Vol. 70 (1), pp. 93-108.

[44] Seim, Katja (2006) "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, Vol. 37 (3), pp. 619-640.

[45] Sorensen, Morten (2007) "How Smart is Smart Money: A Two-Sided Matching Model of Venture Capital," *Journal of Finance*, 62, pp. 2725-2762

[46] Tamer, Elie (2003) "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, Vol 70(1), pp. 147-167.

[47] Uetake, Kosuke and Yasutora Watanabe (2012) "A Note on Estimation of Two-Sided Matching Models," *Economics Letters*, Vol. 116, pp. 535-537.

[48] Wheelock, C. David, and Paul W. Wilson (2000) "Why Do Banks Disappear? The Determinants of U.S. Bank Failures and Acquisitions," *Review of Economics and Statistics*, Vol. 82(1), pp 127–138.

[49] Yip, George S. (1982) "Diversification entry: Internal development versus acquisition," *Strategic Management Journal*, 3(4), pp. 331-345.

7 Appendices

7.1 Appendix A: Proofs

Lemma 1 *Given the estimation function $\mathcal{N}_j(k) = N$, $\forall k \in \overline{K}$, the functions $\lambda^{\min}(\overline{K}, \mathcal{N})$ and $\lambda^{\max}(\overline{K}, \mathcal{N})$ are weakly decreasing in N. Also, $\lambda(\overline{K}, \mathcal{N})$ is upper-hemicontinuous.*

Proof. If N in the estimation function increases to $N+1$ (denoted by $\mathcal{N}^N$ and $\mathcal{N}^{N+1}$), the payoff from merger and entering without merger weakly decreases, i.e., $\Pi_j(k, N+1) \leq \Pi_j(k, N)$ for $\forall k \in \overline{K}_i \backslash \{o\}$; though the payoff from not entering remains the same at zero $\Pi_j(o, N+1) = \Pi_j(o, N) = 0$. Thus, from the definition of chosen set $\overline{C}_i$, weakly more firms choose not to enter. This results in $\lambda^{\min}(\overline{K}, \mathcal{N}^N) \leq \lambda^{\min}(\overline{K}, \mathcal{N}^{N+1})$ and $\lambda^{\max}(\overline{K}, \mathcal{N}^N) \leq \lambda^{\max}(\overline{K}, \mathcal{N}^{N+1})$. Regarding upper-himicontinuity, we consider mixing with probability q_k, which can take any value with $q_k \in [0, 1]$ and $\sum_{\kappa \in \overline{C}_j(\overline{K}_j, \mathcal{N}_j)} q_\kappa = 1$ in case if the chosen set is not singleton. Thus, the graph is closed and the correspondence λ is upper-hemicontinuous. ■

Proposition 1 *There exists N^*such that $N^* \in \lambda(\overline{K}^*, \mathcal{N}^*)$.*

Proof. Since the set $[0, N_{\mathcal{E}} + N_{\mathcal{I}}] \subset \mathbb{R}_+$ is compact and convex, and λ is upper-hemicontinuous, we can apply Kakutani's fixed point theorem. ■

Theorem 2 *The generalized Gale-Shapley algorithm with externalities globally converges to an $\mathcal{N}$-stable outcome. If the algorithm starts from $(K_{\mathcal{E}}, K_{\mathcal{I}}) = (\mathcal{K}, \varnothing)$, then it converges to the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome, $\overline{K}^{*\mathcal{E}}$. Similarly, if the algorithm starts from $(K_{\mathcal{E}}, K_{\mathcal{I}}) = (\varnothing, \mathcal{K})$, then it converges to the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome, $\overline{K}^{*\mathcal{I}}$.*

Proof. Due to Theorem 1, Conditions 2 and 3 for $\mathcal{N}$-stability are always satisfied given N. Thus, the remaining task for the proof is checking Condition 1 (the consistency of the estimation). Because function $\lambda^{\min}(K, N)$ is weakly decreasing, $||\lambda^{\min}(K, N) - N||$ is quasi-convex. Hence, the minimization algorithm globally converges to $N^* = \arg\min ||\lambda^{\min}(\overline{K}^{*\mathcal{E}}(N), N) - N||$. If N^* is an integer (Case A in Figure 2), then we have $\left\| \lambda^{\min}(\overline{K}^{*\mathcal{E}}(N^*), N^*) - N^* \right\| = 0$. Thus, $N^* \in \lambda(\overline{K}^*, N^*)$. If N^* is not an integer (Case B in Figure 2), then we have

$\left\|\lambda^{\min}(\overline{K}^{*\mathcal{E}}(N^*), N^*) - N^*\right\| = q$ for some $q \in [0,1]$. In this case, $N^* \in \lambda(\overline{K}^*, N^*) = [\lambda^{\min}(\overline{K}^{*\mathcal{E}}(N^*), N^*), \lambda^{\max}(\overline{K}^{*\mathcal{E}}(N^*), N^*)]$. Thus, Condition 1 is satisfied at N^*. Finally, using this N^* and Theorem 1, we can obtain $\overline{K}^{*\mathcal{E}} = \overline{K}^{*\mathcal{E}}(N^*)$ by starting from $(K_{\mathcal{E}}, K_{\mathcal{I}}) = (\mathcal{K}, \varnothing)$, and similarly for $\overline{K}^{*\mathcal{I}}$. ■

Corollary 1 *(i) The $\mathcal{E}$-proposing and $\mathcal{I}$-proposing generalized Gale-Shapley algorithms with externalities terminate at the same N^*.*

(ii) An unmatched firm j in an $\mathcal{N}$-stable outcome is also unmatched in any $\mathcal{N}$-stable outcome, and the payoff of firm j is the same in any $\mathcal{N}$-stable outcome.

(iii) In any $\mathcal{N}$-stable outcome $(\overline{K}^*, \mathcal{N}^*)$, the number of mergers is uniquely determined, i.e., $N^*_{merge} = \gamma(\overline{K}^*, \mathcal{N}^*)$ $\forall(\overline{K}^*, \mathcal{N}^*)$.

Proof. (i) From part (iii) of Theorem 1, the $\mathcal{E}$-proposing and $\mathcal{I}$-proposing generalized Gale-Shapley algorithm inner loops yield exactly the same number of operating firms $\lambda(\overline{K}^{*\mathcal{E}}(N), N) = \lambda(\overline{K}^{*\mathcal{I}}(N), N)$ given an estimation N in step 2. Therefore, both $\mathcal{E}$-proposing and $\mathcal{I}$-proposing algorithms satisfy the consistency condition at the same N^*. (ii) From part (iv) of Theorem 1, an unmatched firm in a stable allocation given N is also unmatched in any stable allocation given N. Thus, the same holds under N^*. This implies that the unmatched firms receives the same payoff because the payoff of the outside options do not depend on matchings. (iii) Similarly, the number of operating firms $\lambda(K, N)$ is the same for all stable allocations given N from part (iii) of Theorem 1. This implies that the number of operating firms, N^*, is the same for any $\mathcal{N}$-stable outcome. Also, the argument for (ii) above implies that the number of mergers is $N^*_{merge} = \gamma(\overline{K}^*, \mathcal{N}^*)$ for any $(\overline{K}^*, \mathcal{N}^*)$. ■

Corollary 2 *The $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome is preferred to any other stable outcome by all $e \in \mathcal{E}$ and the least preferred by all $i \in \mathcal{I}$, i.e., for any $\mathcal{N}$-stable outcome $\overline{K}^*$,*

$$\Pi_e(\overline{K}^{*\mathcal{E}}, \mathcal{N}^*) \geq \Pi_e(\overline{K}^*, \mathcal{N}^*) \text{ and } \Pi_i(\overline{K}^*, \mathcal{N}^*) \leq \Pi_i(\overline{K}^{*\mathcal{E}}, \mathcal{N}^*) \tag{11}$$

for the $\mathcal{E}$-optimal $\mathcal{N}$-stable outcome.

Similarly, the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome is preferred to any other stable outcome by all $i \in \mathcal{I}$ and the least preferred by all $e \in \mathcal{E}$, i.e., for any $\mathcal{N}$-stable outcome $\overline{K}^*$,

$$\Pi_i(\overline{K}^{*\mathcal{I}}, \mathcal{N}^*) \geq \Pi_i(\overline{K}^*, \mathcal{N}^*) \text{ and } \Pi_e(\overline{K}^{*\mathcal{I}}, \mathcal{N}^*) \leq \Pi_e(\overline{K}^*, \mathcal{N}^*) \tag{12}$$

for the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome.

Proof. From (ii) of Theorem 1, the inner loop of $\mathcal{E}$-proposing generalized Gale-Shapley algorithm yields the set of stable contracts $\overline{K}^{*\mathcal{E}}(N)$ for a given N, which is unanimously the

most preferred set of contracts among the set of all stable contracts for $\mathcal{E}$ and unanimously the least preferred for $\mathcal{I}$ given estimation N in step 2. This property holds for any given N. Now, Proposition 1 and Corollary 1 imply that the outer loop converges to N^*, and the same statement holds for N^*, which corresponds to an $\mathcal{N}$-stable outcome. The same argument holds for the $\mathcal{I}$-optimal $\mathcal{N}$-stable outcome. ■

7.2 Appendix B: Moment Inequalities and Equalities

We describe the moment inequalities and equalities we use in our estimation. First, we have moment inequalities from mean of payoffs for potential entrants

$$\begin{aligned} E\left[\frac{1}{N_{\mathcal{E}}}\sum_{e\in\mathcal{E}}\left[\Pi_e(\overline{K}^{*\mathcal{E}};\theta)-\Pi_e(\overline{K}^{*DATA};\theta)\right]\right] &\geq 0, \\ E\left[\frac{1}{N_{\mathcal{E}}}\sum_{e\in\mathcal{E}}\left[\Pi_e(\overline{K}^{*DATA};\theta)-\Pi_e(\overline{K}^{*\mathcal{I}};\theta)\right]\right] &\geq 0, \end{aligned}$$

and the same for incumbents as well. Also, we consider the conditional moments for the mean payoffs of the potential entrants,

$$\begin{aligned} E\left[\frac{1}{N_{\mathcal{E}}}\sum_{e\in\mathcal{E}}\left[\Pi_e(\overline{K}^{*\mathcal{E}};\theta)-\Pi_e(\overline{K}^{*DATA};\theta)|X\right]\right] &\geq 0, \\ E\left[\frac{1}{N_{\mathcal{E}}}\sum_{e\in\mathcal{E}}\left[\Pi_e(\overline{K}^{*DATA};\theta)-\Pi_e(\overline{K}^{*\mathcal{I}};\theta)|X\right]\right] &\geq 0, \end{aligned}$$

as well as for the incumbents. In addition to the differences in mean payoff between the observed and upper (and lower) bounds, we consider quantiles as well. Denoting α-quantile among the player $\mathcal{E}$ by $Q_{\alpha,\mathcal{E}}(\cdot)$, we use

$$\begin{aligned} E\left[Q_\alpha\left(\Pi_e(\overline{K}^{*\mathcal{E}};\theta)-\Pi_e(\overline{K}^{*DATA};\theta)\right)\right] &\geq 0, \\ E\left[Q_\alpha\left(\Pi_e(\overline{K}^{*DATA};\theta)-\Pi_e(\overline{K}^{*\mathcal{I}};\theta)\right)\right] &\geq 0, \end{aligned}$$

as well as for the incumbents. The quantiles we use for incumbents are 25%, 50%, and 75%. For potential entrants, we use 95%, and 99% because there is not much information for lower quantiles for potential entrants.

Regarding moment equalities, we have two types of equalities. The first type of equalities are regarding the number of operating firms, i.e.,

$$\begin{aligned} E\left[N_{merge}^{DATA}|X\right] &= E\left[\gamma(\overline{K}^{*\mathcal{E}},N^{DATA};\theta)|X\right]=E\left[\gamma(\overline{K}^{*\mathcal{I}},N^{DATA};\theta)|X\right], \\ E\left[N^{DATA}|X\right] &= E\left[\lambda(\overline{K}^{*\mathcal{E}},N^{DATA};\theta)|X\right]=E\left[\lambda(\overline{K}^{*\mathcal{I}},N^{DATA};\theta)|X\right]. \end{aligned}$$

The second type of equalities are about the payoffs of unmatched firms (from Part (ii) of Corollary 1). We take mean of the payoff of unmatched firms to construct moment equalities. Denoting the set of unmatched firms in each market by UM and the number of the unmatched firms by N_{UM}, we have the following moment equalities;

$$\begin{aligned}
E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*DATA};\theta)\right] &= E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*\mathcal{E}};\theta)\right], \\
E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*DATA};\theta)\right] &= E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*\mathcal{I}};\theta)\right], \\
E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*DATA};\theta)|X\right] &= E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*\mathcal{E}};\theta)|X\right], \\
E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*DATA};\theta)|X\right] &= E\left[\frac{1}{N_{UM}}\sum_{j\in UM}\Pi_j(\overline{K}^{*\mathcal{I}};\theta)|X\right].
\end{aligned}$$

In total, there are 44 moment inequalities and 14 moment equalities.

7.3 Appendix C: Computation of Table 4

The model has 21 parameters, and the confidence set, which we denote as CS, is a 21-dimensional object. As we cannot present a 21-dimensional object in a convenient way, we present the min and max of the CS along each dimension in Table 4. In the following, we explain how we obtained the min and the max of the CS along each dimension.

Following the notation of Andrews and Soares (2010), a parameter value θ is included in CS if $T_n(\theta) \leq \widehat{c}_n(\theta, 1-\alpha)$ where $T_n(\theta)$ is the test statistic and $\widehat{c}_n(\theta, 1-\alpha)$ is the critical value. Denoting the j-th element of θ by θ^j, we report $\underline{\theta}^j = \min\{\theta^j | \theta \in CS\}$ and $\overline{\theta}^j = \max\{\theta^j | \theta \in CS\}$. Though computing CS directly is extremely costly given that the CS has 21 dimensions, we can compute $\underline{\theta}^j$ within manageable time by solving the following constrained optimization problem for each of j-th dimension;

$$\begin{aligned}
&\min_{\theta} \theta^j \\
&s.t.\ T_n(\theta) \leq \widehat{c}_n(\theta, 1-\alpha),
\end{aligned}$$

where $\theta^j$is the j-th element of θ. By maximizing instead of minimizing θ^j, we can obtain $\overline{\theta}^j$. We repeat this for $j = 1, ..., 21$ and report the solutions in Table 4.

7.4 Appendix D: Computation of Counterfactual

Though we cannot directly compute the 21-dimensional confidence set, CS, we can still conduct counterfactual policy experiments by imposing a restriction that the parameter values must satisfy the requirement for being included in CS. In computing the upper and

lower bounds for an outcome of interest y (say, number of operating firms), we solve the following constrained optimization problem similar to the problem in Appendix C;

$$\min_{\theta} y(\theta)$$
$$s.t.\ T_n(\theta) \leq \widehat{c}_n(\theta, 1-\alpha),$$

where the notations are same as in Appendix C. We compute the solution to the problem to have the lower bound for the outcome of interest y, such as mean and median numbers of operating firms. We also solve the corresponding maximization problem to obtain the upper bound for the outcome of interest.

8 Supplementary Materials

8.1 Algorithm for the Case of an Alternative Estimation Function

In this supplementary material, we show that $\mathcal{N}$-stable outcome also exists for another specification of the estimation function $\mathcal{N}_j^N(k) = \begin{cases} N \text{ if } k = j \\ N-1 \text{ if } k = k_{jl}, \text{ or } k = o \end{cases}$, and then provide an algorithm to obtain the stable allocation.

Lemma 2 *Given the estimation function* $\mathcal{N}_j^N(k) = \begin{cases} N \textit{ if } k = j \\ N-1 \textit{ if } k = k_{jl}, \textit{ or } k = o \end{cases}$, $\forall k \in \overline{K}$, $\lambda(\overline{K}, \mathcal{N})$ *is weakly decreasing in* N.

Proof. Consider the estimation function $\mathcal{N}_j^N(k) = \begin{cases} N \text{ if } k = j \\ N-1 \text{ if } k = k_{jl}, \text{ or } k = o \end{cases}$, and increasing N to $N+1$ (from $\mathcal{N}^N$ to $\mathcal{N}^{N+1}$) such that $\mathcal{N}_j^{N+1}(k) = \begin{cases} N+1 \text{ if } k = j \\ N \text{ if } k = k_{jl}, \text{ or } k = o \end{cases}$. We can write the corresponding chosen set for $\mathcal{N}^N$ and $\mathcal{N}^{N+1}$ as

$$\overline{C}_i(\overline{K}_i, \mathcal{N}^N) = \begin{cases} o \text{ if } 0 \geq \max\{\Pi_i(i,N), \max_{k \in K_i} \Pi_i(k, N-1)\} \\ i \text{ if } \Pi_i(i,N) \geq \ \max\{0, \max_{k \in K_i} \Pi_i(k, N-1)\} \\ k_{ei} \text{ if } \Pi_i(k_{ei}, N-1) \geq \max\{0, \Pi_i(i,N), \max_{k \in K_i} \Pi_i(k, N-1)\} \end{cases} \quad \text{and}$$

$$\overline{C}_i(\overline{K}_i, \mathcal{N}^{N+1}) = \begin{cases} o \text{ if } 0 \geq \max\{\Pi_i(i,N+1), \max_{k \in K_i} \Pi_i(k, N)\} \\ i \text{ if } \Pi_i(i,N+1) \geq \ \max\{0, \max_{k \in K_i} \Pi_i(k, N)\} \\ k_{ei} \text{ if } \Pi_i(k_{ei}, N) \geq \max\{0, \Pi_i(i,N+1), \max_{k \in K_i} \Pi_i(k, N)\}. \end{cases}$$

From the property of the payoff we have $\Pi_j(k, N+1) \leq \Pi_j(k, N)$ for $\forall k \in \overline{K}_i \backslash o$. Hence, we have $\max\{\Pi_i(i,N), \max_{k \in K_i} \Pi_i(k, N-1)\} \geq \max\{\Pi_i(i,N+1), \max_{k \in K_i} \Pi_i(k, N)\}$, i.e., number of firms choosing not to enter increases. This results in $\lambda(\overline{K}, \mathcal{N}^N) \leq \lambda(\overline{K}, \mathcal{N}^{N+1})$. ■

Now let us define the fraction of players who have estimation function $\mathcal{N}_j^N(k)$ as ρ and $\mathcal{N}_j^{N+1}(k)$ as $1-\rho$. If we observe N^*, and if we assume estimation function to have one less

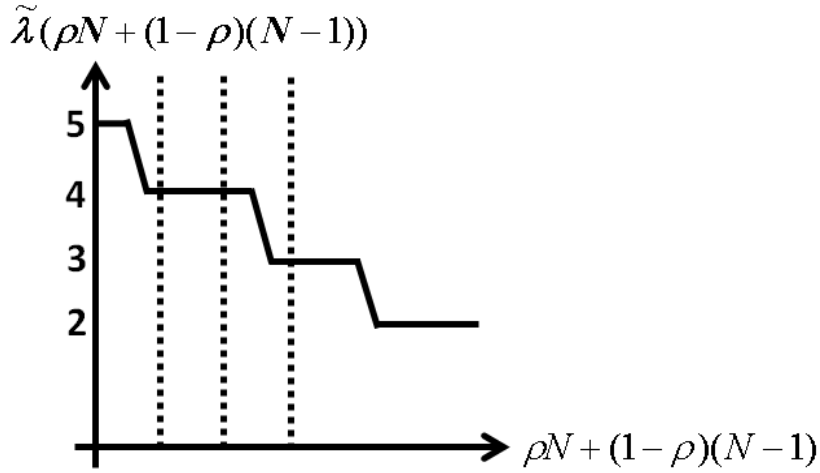

Figure 3: Finding Consistent N with two types of players. Fraction ρ of players have estimation function $\mathcal{N}_j^N$ and Fraction $1-\rho$ of players have estimation function $\mathcal{N}_j^{N-1}$.

number of operating firms if the firm do not enter or enter by merger, then it should be that $\mathcal{N}_j^N(j) = N^*$ and $\mathcal{N}_j^{N+1}(o) = \mathcal{N}_j^{N+1}(k_{jl}) = N^*$. Note that this fraction depends on the choice they make.

Similar to the case of the estimation function discussed in the main text, we consider mixing with probability q in case of indifference. Note that the indifference occurs only to one player generically, and indifference is between entering without merger and not entering. The mixing have a slightly different implication in this case. This is because the change in the choice of the indifferent player affect not only q but also ρ.

Consistency of all players' estimation directly implies that the consistency also holds for the weighted average of the estimations of all players, that is $\rho N + (1-\rho)(N-1)$. Now, we define $\widetilde{\lambda}$, which maps the weighted average of expectation onto the number of operating firms which implicitly solves the optimization behavior of each player given their estimations. Figure 3 is an example of how $\widetilde{\lambda}$ changes with the weighted average of the estimation. Number of operating firms given weighted average of estimation, $\widetilde{\lambda}$, is a function rather than a correspondence because the change in the choice of the indifferent player affect not only q but also ρ. The downward slope part have slope of $-1/(N_{\mathcal{I}}+N_{\mathcal{E}})$, and the effect of change of one player will be smaller as number of players increase. As $\widetilde{\lambda}$ is a function, the existence can be shown using the Brouwer's fixed point theorem.

Lemma 3 *Given the estimation function* $\mathcal{N}_j^N(k) = \begin{cases} N \text{ if } k=j \\ N-1 \text{ if } k=k_{jl}, \text{ or } k=o \end{cases}$, *there exists* N^* *such that* $N^* = \lambda(\overline{K}^*, \mathcal{N}^*)$.

Proof. Because $\widetilde{\lambda}$ is monotonically decreasing and $\widetilde{\lambda}$ is a mapping from $[0, N_{\mathcal{I}} + N_{\mathcal{E}}]$ to $[0, N_{\mathcal{I}}+N_{\mathcal{E}}]$, Brouwer's fixed point theorem implies that there is a fixed point, N^*. Given N^*, we can always find an integer N' and $\rho \in [0, 1)$ such that $N^* = \rho N' + (1-\rho)(N'-1)$. Thus, we have fraction ρ of players to choose not to enter without merger with estimation function $\mathcal{N}_j^{N*}$, i.e., $\mathcal{N}_j^{N*}(j) = N^*$, and fraction $1-\rho$ of the players choosing either not to enter or enter with merger with estimation function $\mathcal{N}_j^{N+1*}$, i.e., $\mathcal{N}_j^{N+1*}(o) = \mathcal{N}_j^{N+1*}(k_{jl}) = N^*$. ■

With this estimation function, taking the model to data have another issue: consistency condition requires some players to have $\mathcal{N}_j^N$ and the others to have $\mathcal{N}_j^{N+1}$. For example, if we observe N in data, the consistency condition needs $\mathcal{N}_j^N(j) = N$ for the player who chooses $\{j\}$, while it requires $\mathcal{N}_j^{N+1}(o) = N$ for the player who chooses $\{o\}$. However, an econometrician cannot know which player is the player who have the estimation function $\mathcal{N}_j^N$ and who are the players with $\mathcal{N}_j^{N+1}$. This unobserved heterogeneity regarding the estimation function of the players can be addressed using EM algorithm in estimation.