MODELLING FIRM MERGERS AS A ROOMMATE PROBLEM

NIKOLAY ANGELOV*

Abstract. We propose a way to model firm mergers using a matching game known as the roommate problem, whereby firms are assumed to make preference rankings of potential merger partners. The position of a firm in another firm's ranking is assumed to be governed by an index, which in turn consists of a deterministic part and of a stochastic one, similar to the latent indices used in standard discrete-choice models. Given all firms' preferences, game-theoretic mechanisms lead to a matching whereby each firm is either self-matched or assigned a merger partner. Assuming that the stochastic part of the preference index is extreme-value distributed, we derive expressions for the probability of a merger between a specific firm pair, and also a log-likelihood function for estimation using firm-specific data. Using a simulation in a setting with groups of three firms involved in roommate games within each group, the model's finite-sample properties are examined.

Key words: firm mergers, roommate game, matching markets, discrete response
JEL classification:

1. Introduction

Mergers and acquisitions have attracted the attention of economists and policymakers for a long time. One important reason is that mergers can have a potentially negative impact on social welfare due to increased market concentration. Thus, it is of interest to measure the effect of mergers on, e.g., consumer prices in the industry of interest.

The present study is concerned with a different aspect, namely the incentives for firms to form mergers. Using common economic terminology, there are several possible motives for two firms to merge. For instance, mergers may allow firms to exploit economies of scale and thus increase their profits. But in addition, increased market power may also lead to higher profits. Since merger incentives reveal firms' beliefs about merger effects, they are also interesting from policymakers' point of view.

The existing literature on merger incentives is abundant, and the present study is connected to at least two research areas: Empirical studies attempting to explain merger

Date: January 25, 2006.
The author is indebted to Per Johansson, Mats Bergman, and John Dagsvik for invaluable advice. Comments from conference participants at the Microdata RTN Meeting, December 2 - 3, 2005 in Copenhagen are appreciated. Financial support from the Research Council of Sweden is gratefully acknowledged.
* *Current address:* Department of Economics, Uppsala University, Box 513, SE-751 20, Uppsala, Sweden. *E-mail address:* nikolay.angelov@nek.uu.se.

probabilities as a function of firm-specific attributes, and more theoretically oriented work from the industrial organization (IO) field. Specifying a reasonable model for firm mergers implies difficult challenges, because of the complex dependence among firms' choices. Firms taking part in a merger formation cannot really make a choice of merging with another firm. At most, the steering board may want to participate in a merger, and take appropriate actions. That may or may not lead to an actual merger, depending of course on what the potential partner's steering board and shareholders want, but also on actions taken by other firms.

Recent theoretical studies of endogenous merger theory typically address this complexity using game theoretic tools.[1] But most of the empirical studies of the motives behind mergers do not employ a strict economic model of mergers. Thus there exists a gap between economic theory and empirical studies.

The present paper addresses this gap by providing a decision framework at individual firm level, that can be used directly in applied work. We argue that firm mergers can be seen as an instance of the roommate game. Furthermore, we aim at making the parameters of the merger model estimable. The analysis is intended to be general, so as to allow for an analysis of pure roommate games, or other situations that can be translated into a roommate problem.

In the IO literature, merger behaviour is often analyzed using game theory as a tool. The present study is most closely related to endogenous merger theory, which Horn and Persson (2001) is an example of. Horn and Persson introduce a game-theoretic model of mergers, not very unlike the one proposed in section 2 of the present paper. At the outset, a number of firms are allowed to form mergers with each other, with a resulting market structure. Each market structure results in a certain profit for the involved firms. If we analyze three firms, it might be that if firms one and two merge, they get a profit of a, while firm three (which has not merged) gets b; or, if firms one and three merge, they get c while firm two gets d, etc. The firms within a merger are then free to bargain on how to share the profit. As will be seen in the sequel, in the firm-merger model we propose, an individual firm cannot be forced to join a merger. In Horn and Persson (2001), however, a market structure is chosen if a so called decisive group has higher combined profit in it than in the alternative market structure. This in turn means that a firm could have lower profit in the chosen market structure than in the alternative one. Thus, an individual firm may be forced by the decisive group to accept a market structure.

[1]The articles by Kamien and Zang (1990), Fridolfsson and Stennek (2005, 2004) Horn and Persson (2001) fall within the field of endogenous firm mergers. Gowrisankaran (1999) is one of few studies that present a dynamic model of endogenous mergers. A typical dynamic effect might be that market concentration increases due to a merger, but may be counterbalanced by increased possibilities of entry by new firms, and thus the welfare effect for society may not be negative. Gowrisankaran (1999) bases his model on the work by Ericson and Pakes (1995).

The merger formation mechanism in Fridolfsson and Stennek (2005, 2004) is of a different kind. Firms take turns in submitting merger proposals to their competitors, which are either accepted or rejected. The model considers a concentrated market with three firms. Once a firm has accepted a proposal, the bargaining ends; otherwise, a new bidding round starts at a random point in time. Each firm has a strategy, describing whether and how much to bid and a reservation price at which to accept an offer from another firm.

The theoretical studies above provide interesting insights into the motives for mergers, and also explain empirical observations made by other studies. For instance, Fridolfsson and Stennek (2004) give a reconciliation between theory and the paradox of why mergers often lead to a rise in consumer prices, which in turn can be assumed to lead to an increase in competitors' profits, while competitors' share prices fall. Nevertheless, it is clear that the models are theoretical, in the sense that model parameters are not easily estimable. An econometrician cannot readily use them in her or his econometric specification. At most, he or she can refer to theoretical results when trying to explain the empirical ones. As will become clear in the sequel, this is the main difference between the present and earlier studies: Our merger game is explicitly converted into an estimable econometric model, giving a clear-cut link between economic theory and empirical estimation. At a later stage in the paper we are forced to restrict the generality of the analysis due to the complexity of the model. We begin with a general economic model of firm mergers, but when the estimation technique is outlined the analysis is limited to oligopoly situations with groups of three firms that take part in different merger games. This limitation is discussed at the end of the paper.

The roommate game is well-studied in the game-theoretic literature. It is a part of the literature on matching markets. The book by Roth and Sotomayor (1990) treats two-sided matching in detail, but the roommate game does not fall within this sub-category. Two-sided matching is not suitable for analyzing firm mergers, because it deals with situations where the group of agents can be naturally divided into two subgroups. In terms of mergers, this would mean that the firms are exogenously divided into, say, acquirers and targets.

A well-known example a two-sided matching is the marriage game. **?** develops a general framework suited for analysis of two-sided matching markets. Under specific distributional assumptions regarding the agents' preferences, he derives a functional relationship between the number of realized matches of a certain type (i.e., corresponding to a certain combination of attributes of agents from each type) and the number of agents with specific attributes from each type. Using this relationship, parameters from the preferences of the agents can be estimated. The framework is applied to a two-sex marriage model in **?**, where men and women are grouped with respect to age. Unfortunately, Dagsvik's framework cannot be applied to firm mergers, because it is only applicable to two-sided matching markets.

The roommate game does not assume anything about the role of the agents beforehand. A simple version is as follows:[2] Assume that an even number of persons, say k, wish to divide up into pairs of roommates to share $k/2$ rooms. Each person ranks the remaining $k-1$ in descending order, beginning with the person most preferred to share room with. A set of pairings, also called a matching, is called stable if there are not two persons – currently not sharing room – who prefer each other to their actual roommates. In contrast to the marriage game, there are examples of preferences for which no stable set of pairings exists.

A number of theoretical articles consider the roommate problem, and the most relevant of those will be mentioned in Section 2. Most of the previous work deals with the issue of stability, i.e., with questions such as whether there exists a stable matching for a specific roommate problem (to be defined in the sequel), and if so, how to reach such a matching. The present paper instead focuses on introducing an econometric model, stemming from the roommate game. To make estimation possible, we introduce randomness in the preference ordering. The resulting econometric specification is much like a discrete choice model, the main difference being that no agent can choose a roommate the way that individuals can make choices in a discrete choice model. Instead, an agent makes a ranking, and then a game-theoretic mechanism leads to a set of pairings.

The next section states the formal economic model of mergers. Sections 3 and 4 deal with estimation issues. We conclude our analysis in section 5, and figures are left for the appendix.

2. The merger game

We need to make some generalizations to the simple roommate game described above in order to use it for modelling mergers. Firms can (and often do) stand on their own, i.e., they do not have to merge with other firms. Thus n, the number of firms, does not have to be even, and more importantly, the number of mergers is not known a priori, but depends on the firms' preferences for one another. The latter implies that not only the merger decisions, but also the number of mergers, is endogenous. Thus we assume that the number of mergers lies between zero and $n/2$.

The rest of this section begins with a formal model of a merger game. At this first stage, we assume that the preference ordering of each firm over the others is given. But the interest of this paper is to go deeper than that; we want to know how firm characteristics

[2]Gale and Shapley (1962, Example 3 on p. 12) is an early reference.

influence the preference orderings, and thus the merger decisions. Therefore, stage two provides a functional relationship between firm attributes and merger formation.

Assume a finite set of firms denoted by $F = \{f_1, f_2, \dots, f_n\}$. Firm f_i's preference ordering over the other firms is denoted by W_i and might for example be of the form

$$W_i = \{f_2, f_1, f_i, f_6, \dots, f_n\}, \tag{1}$$

implying that firm f_i's first choice is, if possible, to merge with firm f_2. If that is impossible, its second choice is to merge with firm f_1, and if that also is unattainable, the firm prefers to continue operating on its own. If f_i strictly prefers to merge with f_k rather than with f_m, we will sometimes write $f_k \succ_{f_i} f_m$ when we do not need to show the complete ordering. At this stage we do not assume strict preferences; that is, we might have week preferences denoted by $f_k \succsim_{f_i} f_m$. This notation means that f_i finds f_k at least as good as f_m. We do, however, assume that preferences are rational, i.e., that they are transitive and form a complete ordering.

Now we make still one assumption, namely that no firm can be forced into a merger, thereby excluding the possibility of hostile takeovers. This assumption implies that the only part of (1) that matters is

$$W_i = \{f_2, f_1, f_i\},$$

because firm f_i prefers to operate alone rather than merge with any firms in the set $\{f_6, \dots, f_n\}$. Another way of stating this is to say that none of the firms in $\{f_6, \dots, f_n\}$ is acceptable to f_i. The collected preference orderings of all firms are called the preference profile and denoted by $\mathcal{W} = \{W_1, W_2, \dots, W_n\}$.

At the outset of the game, all n firms operate on their own, and at this stage we assume that their preference orderings are known; later this assumption will be dropped. The purpose of the game is a set of pairings, given the preference profile. Formally, this can be expressed in terms of a matching, which is defined below.

Definition 1 (Matching)**.** A matching μ implies that each firm is either self-matched (i.e., continues to operate on its own), or is matched to at most one merge partner. A stable matching is one where all matches are individually acceptable (i.e., none of the matched firms prefers to be self-matched rather than its current matching), and in addition, no pair of firms prefer to be matched to each other rather than according to the prevailing matching.

An example might be useful at this stage. Set $n = 3$ and assume that the preference profile is $\mathcal{W} = \{\{f_3, f_1, f_2\}, \{f_1, f_3, f_2\}, \{f_2, f_3, f_1\}\}$. Consider the matching μ defined by a merger between f_1 and f_3, while f_2 continues on its own. According to the preference profile, f_1 is satisfied, since it merges with its most preferred choice. But f_3 prefers f_2 to f_1, and since f_2 prefers f_3 to not merging at all, f_2 and f_3 would gain if they merged with each other. Therefore, the matching is not stable. Consider instead the matching denoted by μ' where f_2 merges with f_3 and f_1 is self-matched. Now f_1 would benefit from merging with f_3, but f_3 prefers its current matching. Furthermore, f_1 and f_2 do not want to form a merger, since they are not acceptable to each other. Consequently, μ' is a stable matching.

As shown in Gale and Shapley (1962), it is easy to construct examples of preference profiles leading to non-existence of a stable matching in the roommate game. Below we will discuss conditions for the existence of a stable matching, but before continuing, we need to introduce the concept of blocking.

Definition 2 (Blocking)**.** A matching μ is blocked by a pair of firms $\{f_i, f_j\} \subseteq F$ if f_i and f_j both prefer each other to the firms they are matched to in μ. This is valid also for $i = j$, i.e., an individual firm f_i can block μ if it does not accept its current merging partner.

The notion of blocking is closely related to our assumption concerning hostile takeovers – if blocking is permitted, there can be no hostile takeovers. Stated in words, blocking implies that each individual firm is allowed to choose the best attainable matching for a given preference profile.

The earliest general treatment of stability in the roommate game can be found in Irving (1985). Given a preference profile, the algorithm proposed by Irving determines whether there exists a stable matching, and if so, finds such a matching. But since the algorithm is defined in terms of a computer program, it is hard to summarize, and what is more important, it does not give any insight regarding the general conditions on preferences that lead to a stable matching. Thus, for a given preference profile that does not lead to a stable matching, one would have to trace Irving's algorithm step by step in order to find out the reason for non-stability.

Later studies provide conditions for stable matchings that are more easily interpreted. There are two sets of results in the literature concerning stability in the roommate game, distinguished by whether preferences are assumed to be weak or strict. The paper by Tan (1991) provides a necessary and sufficient condition for the existence of a stable matching in the case of strict preferences, which is stated in terms of a preference restriction. This condition is rather technical, and since it will not be used in the present paper, the interested reader is referred to the article. We will instead focus on the more general treatment in Chung (2000) which allows for weak preferences, and in addition supplies a condition which is somewhat less tedious to state, and will therefore prove more suitable for our needs. Chung identifies a sufficient condition for the existence of stable roommate matchings, called the no odd rings condition. Before stating it, we need one more definition:

Definition 3 (Odd ring)**.** A ring is an ordered subset of firms $\{f_1, f_2, \ldots, f_k\}$, $k \geq 3$, such that for $1 \leq i \leq k$,

$$f_{i+1} \succ_{f_i} f_{i-1} \succsim_{f_i} f_i, \quad \text{for odd } i \text{, and}$$
$$f_{i+1} \succsim_{f_i} f_{i-1} \succsim_{f_i} f_i \quad \text{for even } i.$$

The subscript is taken modulo k, which is explained by the following example: For $k = 3$ and $i = 1$, f_{i-1} is read f_3, because $3 \equiv i - 1 \pmod 3 = 0 \pmod 3$. An odd ring is a ring such that k is odd, and a strict odd ring is an odd ring in the special case with strict preferences. (Note that in the latter case, distinguishing between odd and even i is not needed.)

This seemingly abstract definition will be treated later on in the paper, and for the special case when $n = 3$ we will explicitly show all possible odd rings. The main result in Chung (2000) is that if there are no odd rings in $\mathcal{W}$, there exist stable roommate matchings. Furthermore, a stable matching can be reached using a simple algorithm, first described in Roth and Vande Vate (1990). In the present setting involving n firms with the preference profile $\mathcal{W}$, this mechanism, which we call the Roth-Vande Vate algorithm, has the following essence: It starts with an arbitrary matching denoted by μ_1. If μ_1 is stable, there will be no blocking pairs, and the algorithm stops. If it is unstable, there will be at least one blocking pair. In that situation we form a new matching denoted by μ_2, where a randomly chosen pair from the ones that block μ_1 is allowed to merge, and their partners under μ_1 are self-matched under μ_2. Now, if μ_2 is stable, the algorithm stops, and if it is unstable, we form μ_3, etc. Chung shows that if there are no odd rings in $\mathcal{W}$, the sequence $\{\mu_i\}_{i=1}^{\infty}$ will converge to a stable roommate matching with probability one.

In general, Chung's no odd rings condition is sufficient, but not necessary.[3] But he shows that when preferences are weak and there exists a stable matching despite the existence of an odd ring, the Roth-Vande Vate algorithm cannot be used to reach the stable matching. On the other hand, if a stable matching exists (with or without the existence of odd rings), and in addition preferences are strict, we may use the result from Theorem 1 in Diamantoudi, Miyagawa, and Xue (2004), stating that the Roth-Vande Vate algorithm can be used.

Is the above algorithm a plausible description of the actual process of merger formation in an industry? Among many other things, the answer depends on whether or not we allow hostile takeovers. If we define a hostile takeover as any firm being forced into a merger against its will, and allow it, no blocking can take place. If we, on the contrary,

[3]For an example of preferences with an odd ring where a stable matching exists, see Chung (2000, Example 2)

exclude the possibility of hostile takeovers, the algorithm seems rather attractive: It can be seen as a way for firms to choose freely whether to merge or not, and in addition, which partner to choose. For the sake of simplicity we therefore assume in the present paper that hostile takeovers do not take place. Consider for instance the situation where f_i is facing bankruptcy and f_j approaches it with a takeover proposition. Even though f_i may be considered forced to accept, in this paper we see it another way: the owners of f_i use their free will to choose the firm being taken over, rather than for it to go bankrupt.

3. A discrete-response framework

Up to this point, we have taken the firms' preference profile as given, which is an unrealistic assumption for an empirical application. It is true that in theory, one could think of survey data with firms' rankings of each other. Using those, a researcher might then analyze possible stable matchings. But such data are bound to be difficult – if not impossible – to obtain. Firstly, mergers are often kept secret until the day of the announcement of a merger proposal, and it is not likely that a researcher would be able to obtain advance information. After a merger proposal is announced, the steering board of a firm would probably find it unappropriate to give information about the ranking of its other potential merger candidates, because such information would probably affect the outcome of the ongoing merger or merger negotiations. Secondly, if the firms of interest are limited liability companies, for practical purposes such a survey would have to be addressed toward the steering boards, and not toward the shareholders. But it is the latter that in fact take the final decision, and it is not unusual that shareholders do not follow the steering board's advice. Consequently, even if a ranking in theory could be obtained, it is not likely that it is the correct one.

This section provides an empirical framework based on a specific distributional assumption regarding the preference profile. We will consider all possible outcomes of a merger game and derive their probabilities. A log-likelihood function and a corresponding maximum-likelihood estimator of the population parameter vector is derived and its finite-sample properties examined, and as will become clear in the sequel, our approach has many similarities with discrete-choice models. The data we have in mind is assumed to consist of several groups of firms, where a roommate merger game takes place among the firms within each group. Along the way, the analysis will be restricted to a maximum of three firms in each group due to the complexity of the resulting probability expressions. This limitation is further discussed at the end of the paper.

To this end, assume that we are dealing with a one-shot game in which n firms participate. The main idea is that each firm f_i, $i = 1, 2, \ldots, n$, observes every other firm's attributes, and uses those (possibly relating them to its own attributes) to make a preference ordering of all firms, including f_i itself. Collect f_i's attributes in the vector $\boldsymbol{x}_i$ and assume the following:

Assumption 1. The position of f_j in f_i's preference ordering solely depends on the index $\Pi_{ij} = g(\boldsymbol{x}_i, \boldsymbol{x}_j) + \varepsilon_{ij}$, where we let $j = 1, 2, \ldots, n$. The function $g(\cdot)$ is deterministic and assumed to be known up to an unknown parameter vector and ε_{ij} are random error terms. Each ε_{ij} is i.i.d. extreme value with probability density function $f_{\varepsilon_{ij}}(\varepsilon_{ij}) = e^{-\varepsilon_{ij}} e^{-e^{-\varepsilon_{ij}}}$.

Note that this assumption ensures strict preferences, since the indices have continuous cumulative distribution functions. The index might for instance represent the profit to f_i's owners from a merger with f_j, or it might be seen as an aggregated utility index that the owners of f_i get from merging f_j. Both parts of the index (and thus the ranking) are assumed to be known by the firm, but only the deterministic part (up to a set of parameters) is known to the econometrician. From the perspective of the researcher, the variables ε_{ij} are treated as random because they influence the index, but are known only to the firm. Thus ε_{ij} is simply defined so as to capture the difference between the true value of the index and the part that is known to the researcher. If $\Pi_{ij} > \Pi_{ik}$, then f_j is ranked higher than f_k in W_i, and thus, there is a one-to-one correspondence from Π_{ij} to $W_i \ \forall \ j$. Note that by construction, the level of the index does not matter for the ranking, solely the order does. Moreover, f_i does not have to place f_j at the same position in its preference ordering, as f_j places f_i.

To get a feeling for the construction of the index, consider a hypothetical example of a matching game among firms of different sizes with locations spread over a whole continent. Assume that in their search of a merging partner, firms' rankings of each other are influenced by geographical distance and by differences in firm size. Then if $\boldsymbol{x}_k = (s_k, l_k)$ for $k = i, j$, where s is size and l is location, we might have $g(\boldsymbol{x}_i, \boldsymbol{x}_j) = g(s_i - s_j, l_i - l_j)$.

Let $\mathcal{I}_{ij} = 1$ if f_i and f_j merge, and zero otherwise, and consider the probability of a merger between the two firms, denoted by $\mathsf{P}_{ij} \equiv \mathsf{P}(\mathcal{I}_{ij} = 1)$. For a given number of firms, we are interested in deriving the probability of a merger between any specific firm pair, and the probability that the all firms are self-matched. In the sequel, we restrict the analysis to the case when $n = 3$. This restriction is imposed simply because of the complexity of the resulting expressions. It can be seen as a limitation to oligopoly situations with groups of three firms that take part in different merger games, and will be further discussed in the last section of this paper. Generalization of the results to a larger number of firms is in theory straightforward, but since no general formulas are available, it is bound to be time-consuming.

Consider first the possible outcomes resulting from a game with three firms. If f_1 merges with f_2, then f_3 must necessarily be self matched, and so this is one possible matching. The same applies to a merger between f_1 and f_3, and f_2 and f_3. Next we have the possibility of all firms being self-matched, and finally, the possibility of preferences leading to no stable matching. Thus, the sample space of a game with three firms consists of five mutually exclusive outcomes and is denoted by $\mathcal{S} = \{\{\mathcal{I}_{11} = \mathcal{I}_{22} = \mathcal{I}_{33} = 1\}, \{\mathcal{I}_{33} = \mathcal{I}_{12} = 1\}, \{\mathcal{I}_{22} = \mathcal{I}_{13} = 1\}, \{\mathcal{I}_{11} = \mathcal{I}_{23} = 1\}, \{\mathcal{I}_{11} = \mathcal{I}_{22} = \mathcal{I}_{33} = \mathcal{I}_{12} = \mathcal{I}_{13} = \mathcal{I}_{23} = 0\}\} \equiv \{A, B, C, D, E\}$. It is easy to check that we must have $\mathsf{P}(A)+\mathsf{P}(B)+\mathsf{P}(C)+\mathsf{P}(D)+\mathsf{P}(E) = 1$.

For reasons which will soon become clear, we begin by considering a specific event, rather than an outcome. The probability of a merger between f_1 and f_2 is given below. For conciseness, below each individual probability is its corresponding short-hand denotation $p_1, p_2, \ldots, p_{11}$.

$$\begin{aligned}
\mathsf{P}(\mathcal{I}_{12} = 1) = \\
&\underbrace{\mathsf{P}(\Pi_{12} > \max\{\Pi_{11}, \Pi_{13}\})}_{p_1} \times \underbrace{\mathsf{P}(\Pi_{21} > \max\{\Pi_{22}, \Pi_{23}\})}_{p_2} \\
&+ \underbrace{\mathsf{P}(\Pi_{13} > \Pi_{12} > \Pi_{11})}_{p_3} \times \underbrace{\mathsf{P}(\Pi_{21} > \max\{\Pi_{22}, \Pi_{23}\})}_{p_4} \times \underbrace{\mathsf{P}(\Pi_{33} > \Pi_{31})}_{p_5} \\
&+ \underbrace{\mathsf{P}(\Pi_{23} > \Pi_{21} > \Pi_{22})}_{p_6} \times \underbrace{\mathsf{P}(\Pi_{12} > \max\{\Pi_{11}, \Pi_{13}\})}_{p_7} \times \underbrace{\mathsf{P}(\Pi_{33} > \Pi_{32})}_{p_8} \\
&+ \underbrace{\mathsf{P}(\Pi_{13} > \Pi_{12} > \Pi_{11})}_{p_9} \times \underbrace{\mathsf{P}(\Pi_{23} > \Pi_{21} > \Pi_{22})}_{p_{10}} \times \underbrace{\mathsf{P}(\Pi_{33} > \max\{\Pi_{31}, \Pi_{32}\})}_{p_{11}}.
\end{aligned} \tag{2}$$

This expression looks disorganized, but is in fact rather intuitive. (For conciseness, when we are considering two specific firms, we will sometimes refer to them as "he" and "she".) It is a sum of four terms, with the first term covering the probability of f_1 giving f_2 highest rank in her preference ordering, at the same time as f_2 ranks f_1 at the top of his. If this is the case, both firms will clearly block any other matching, and we will end up with a merger between them. The second term deals with the case when f_2 acceptable to f_1, but is ranked lower than f_3, at the same time as f_1 is the most preferred in f_2's ordering. In this case, $\mathcal{I}_{12} = 1$ if and only if f_1 is not acceptable to f_3. The third term is analogous to the second, the roles of f_1 and f_2 being reversed. Finally, the last term covers the case when f_1 and f_2 are acceptable to each other, but each prefers f_3 the most. Here, f_1 and f_2 will still merge if none of them is acceptable to f_3.

Given the distributional properties of Π_{ij} given in Assumption 1, we will derive a closed-form formula for (2). First, notice that $p_4 = p_2$, $p_7 = p_1$, $p_9 = p_3$, and $p_{10} = p_6$, and thus we only need to attend to p_1, p_2, p_3, p_5, p_6, p_8, and p_{11}. Consider p_1, p_2, p_5, p_8, and p_{11}, which can all be written on any of the forms $\mathsf{P}(\Pi_{ij} > \max\{\Pi_{ik}, \Pi_{il}\})$ or $\mathsf{P}(\Pi_{ij} > \Pi_{ik})$. A well-known result is readily applicable to this case, namely the formula for logit choice probabilities which is derived in e.g., McFadden (1974) using the same distributional assumptions as here. Notice that we are not referring to the logit model, but merely to the expression for logit choice probabilities. The formula is defined for an agent i making a choice among J alternatives. Each alternative $j = 1, 2, \dots, J$ gives to i the utility Π_{ij}, which is according to Assumption 1. A utility-maximizing agent i will choose alternative j if $\Pi_{ij} > \Pi_{ik}\ \forall\ k \neq j$, and the probability for this, also called the logit choice probability, is given by

$$\mathsf{P}(\Pi_{ij} > \Pi_{ik}\ \forall\ k \neq j) = \frac{e^{V_{ij}}}{\sum_{m=1}^{J} e^{V_{im}}},$$

where $V_{ij} \equiv g(\boldsymbol{x}_i, \boldsymbol{x}_j)$. This formula is directly applicable to the five expressions mentioned above and using it results in

$$\begin{aligned} p_1 &= \frac{e^{V_{12}}}{e^{V_{11}} + e^{V_{12}} + e^{V_{13}}} \qquad p_2 = \frac{e^{V_{21}}}{e^{V_{21}} + e^{V_{22}} + e^{V_{23}}} \qquad p_5 = \frac{e^{V_{33}}}{e^{V_{31}} + e^{V_{33}}} \\ p_8 &= \frac{e^{V_{33}}}{e^{V_{32}} + e^{V_{33}}} \qquad p_{11} = \frac{e^{V_{33}}}{e^{V_{31}} + e^{V_{32}} + e^{V_{33}}}. \end{aligned} \tag{3}$$

The formula for logit choice probabilities cannot be applied to the remaining probabilities p_3 and p_6. Nevertheless, there are results in the literature that can be used. McFadden (1984, p. 1414) presents a formula for the probability of an observed ranking of m of the alternatives in a choice set of a logit model, which in our setting can be written as

$$\mathsf{P}(\Pi_{i1} > \Pi_{i2} > \cdots > \Pi_{im}) = \frac{e^{V_{i1}}}{\sum_{k=1}^{m} e^{V_{ik}}} \frac{e^{V_{i2}}}{\sum_{k=2}^{m} e^{V_{ik}}} \cdots \frac{e^{V_{i(m-1)}}}{\sum_{k=m-1}^{m} e^{V_{ik}}}. \tag{4}$$

Using this gives the following analytical expressions for p_3 and p_6:

$$p_3 = \frac{e^{V_{13}}}{e^{V_{11}} + e^{V_{12}} + e^{V_{13}}} \frac{e^{V_{12}}}{e^{V_{11}} + e^{V_{12}}} \qquad p_6 = \frac{e^{V_{23}}}{e^{V_{21}} + e^{V_{22}} + e^{V_{23}}} \frac{e^{V_{21}}}{e^{V_{21}} + e^{V_{22}}} \tag{5}$$

The formulas in (3) and (5) inserted in (2) result in an analytical expression for the probability of a merger between f_1 and f_2, given by

$$\begin{aligned} \mathsf{P}_{12} &= p_1 p_2 + p_3 p_4 p_5 + p_6 p_7 p_8 + p_9 p_{10} p_{11} \\ &= \frac{e^{V_{12}}}{\sum_{j=1}^{3} e^{V_{1j}}} \frac{e^{V_{21}}}{\sum_{j=1}^{3} e^{V_{2j}}} + \frac{e^{V_{13}}}{\sum_{j=1}^{3} e^{V_{1j}}} \frac{e^{V_{12}}}{(e^{V_{11}} + e^{V_{12}})} \frac{e^{V_{21}}}{\sum_{j=1}^{3} e^{V_{2j}}} \frac{e^{V_{33}}}{(e^{V_{31}} + e^{V_{33}})} \\ &\quad + \frac{e^{V_{23}}}{\sum_{j=1}^{3} e^{V_{2j}}} \frac{e^{V_{21}}}{(e^{V_{21}} + e^{V_{22}})} \frac{e^{V_{12}}}{\sum_{j=1}^{3} e^{V_{1j}}} \frac{e^{V_{33}}}{(e^{V_{32}} + e^{V_{33}})} \\ &\quad + \frac{e^{V_{13}}}{\sum_{j=1}^{3} e^{V_{1j}}} \frac{e^{V_{12}}}{(e^{V_{11}} + e^{V_{12}})} \frac{e^{V_{23}}}{\sum_{j=1}^{3} e^{V_{2j}}} \frac{e^{V_{21}}}{(e^{V_{21}} + e^{V_{22}})} \frac{e^{V_{33}}}{\sum_{j=1}^{3} e^{V_{3j}}}. \end{aligned}$$

Simplifying this results in

$$\mathsf{P}_{12} = \frac{e^{V_{12}+V_{21}}}{\sum_{j=1}^{3} e^{V_{1j}} \sum_{j=1}^{3} e^{V_{2j}}} \Bigg(1 + \frac{e^{V_{13}+V_{33}}}{(e^{V_{11}} + e^{V_{12}})(e^{V_{31}} + e^{V_{33}})} + \frac{e^{V_{23}+V_{33}}}{(e^{V_{21}} + e^{V_{22}})(e^{V_{32}} + e^{V_{33}})} + \frac{e^{\sum_{k=1}^{3} V_{k3}}}{(e^{V_{11}} + e^{V_{12}})(e^{V_{21}} + e^{V_{22}}) \sum_{j=1}^{3} e^{V_{3j}}} \Bigg). \tag{6}$$

The probabilities of a merger between f_1 and f_3, and f_2 and f_3 can be written in a similar way. It is simply a matter of changing the indices in the expression above.

Next, consider the probability of the non-existence of a stable matching. Remember that in general, the condition of no odd rings is sufficient for the existence of a stable matching. But for the special case of three firms, it is also necessary, i.e., a stable matching cannot exist when an odd ring exists. This can be easily checked using Definition 3, as shown below. A strict odd ring exists if the preference profile is either $\mathcal{W}_1 = \{\{f_2, f_3, f_1\}, \{f_3, f_1, f_2\}, \{f_1, f_2, f_3\}\}$ or $\mathcal{W}_2 = \{\{f_3, f_2, f_1\}, \{f_1, f_3, f_2\}, \{f_2, f_1, f_3\}\}$. To see why, consider the ordered set $\{f_1, f_2, f_3\}$. Using Definition 3, we can derive a preference profile with a strict odd ring, namely

$$\begin{aligned} &f_2 \succ_{f_1} f_3 \succ_{f_1} f_1, \\ &f_3 \succ_{f_2} f_1 \succ_{f_2} f_2, \text{ and} \\ &f_1 \succ_{f_3} f_2 \succ_{f_3} f_3, \end{aligned}$$

which is just another way of writing $\mathcal{W}_1$.

In order to see where $\mathcal{W}_2$ comes from, make the following name changes: $f_1 \to c_2$, $f_2 \to c_1$, and $f_3 \to c_3$. Obviously, such a name change should not alter the preferences in any real way – for instance, if f_1 places f_2 at the top of her preferences, c_2 should place c_1 on top of hers. Thus, using the new names, $\mathcal{W}_1$ can be written as

$$\begin{aligned} &c_3 \succ_{c_1} c_2 \succ_{c_1} c_1, \\ &c_1 \succ_{c_2} c_3 \succ_{c_2} c_2, \text{ and} \\ &c_2 \succ_{c_3} c_1 \succ_{c_3} c_3. \end{aligned}$$

Changing names from c_i to f_i for $i = 1, 2, 3$ gives us the preference profile $\mathcal{W}_2$.

If we start out from $\mathcal{W}_1$, but instead make the name changes $f_1 \to c_1$, $f_2 \to c_3$, and $f_3 \to c_2$, we still end up with $\mathcal{W}_2$. Since the same applies to the name changes $f_1 \to c_3$, $f_2 \to c_2$, and $f_3 \to c_1$, we can conclude that the only two preference profiles giving rise to a strict odd ring are $\mathcal{W}_1$ and $\mathcal{W}_2$.

It is easily checked that the preference profiles $\mathcal{W}_1$ and $\mathcal{W}_2$ never lead to a stable matching. This can be done by considering each of the four possible matchings in $\mathcal{S}$ for the preference profiles $\mathcal{W}_1$ or $\mathcal{W}_2$. For each matching and preference profile, there will always be a blocking pair, which is another way of saying that there does not exist any stable matching. For instance, consider the matching $\mu_1 = \{\mathcal{I}_{33} = \mathcal{I}_{12} = 1\}$ with preference profile $\mathcal{W}_1$. Firm f_1 is matched to its most-wanted partner, and has no incentives to block μ_1, but f_2 prefers to be matched to f_3, and f_3 prefers to be matched to f_2, resulting in the blocking pair $\{f_2, f_3\}$. Satisfying the blocking pair leads to $\mu_2 = \{\mathcal{I}_{11} = \mathcal{I}_{23} = 1\}$. But μ_2 is blocked by $\{f_1, f_3\}$, because f_1 ranks f_3 higher than f_2, and f_3 ranks f_1 higher than f_2. Satisfying the blocking pair leads to $\mu_3 = \{\mathcal{I}_{22} = \mathcal{I}_{13} = 1\}$, which in turn is blocked by $\{f_1, f_2\}$, because f_1 ranks f_2 higher than f_3 and f_2 ranks f_1 higher than itself. But satisfying $\{f_1, f_2\}$ leads to μ_1, which starts the blocking procedure all over again. What about the matching $\mu_4 = \{\mathcal{I}_{11} = \mathcal{I}_{22} = \mathcal{I}_{33} = 1\}$? There are several blocking pairs, each of them necessarily leading to one of the matchings μ_1–μ_3 and thus to the unstable blocking procedure described above. For instance, $\{f_1, f_2\}$ is a blocking pair because each of the firms rank one another higher than being self-matched. But satisfying them leads to μ_1.

In a similar way, it can be shown that the preference profile $\mathcal{W}_2$ cannot lead to a stable matching, and thus for the case of three firms, the no odd rings condition is both a sufficient and a necessary one. Therefore the probability of no stable matchings is equal to the probability of a strict odd ring, i.e., of observing $\mathcal{W}_1$ or $\mathcal{W}_1$. In terms of preference index comparisons, this can be written as

$$\begin{aligned}\mathsf{P}_E = {} & \mathsf{P}(\Pi_{12} > \underset{p_{12}}{\Pi_{13}} > \Pi_{11})\mathsf{P}(\Pi_{23} > \underset{p_{13}}{\Pi_{21}} > \Pi_{22})\mathsf{P}(\Pi_{31} > \underset{p_{14}}{\Pi_{32}} > \Pi_{33}) \\ & + \mathsf{P}(\Pi_{13} > \underset{p_{15}}{\Pi_{12}} > \Pi_{11})\mathsf{P}(\Pi_{21} > \underset{p_{16}}{\Pi_{23}} > \Pi_{22})\mathsf{P}(\Pi_{32} > \underset{p_{17}}{\Pi_{31}} > \Pi_{33}).\end{aligned} \tag{7}$$

Applying (4) from p. 11 on this results in the following analytical expression for the probability of an odd ring:

$$\begin{aligned}\mathsf{P}_E &= p_{12}p_{13}p_{14} + p_{15}p_{16}p_{17} \\ &= \frac{e^{V_{12}}e^{V_{13}}}{\sum_{j=1}^{3} e^{V_{1j}}(e^{V_{13}} + e^{V_{11}})} \frac{e^{V_{23}}e^{V_{21}}}{\sum_{j=1}^{3} e^{V_{2j}}(e^{V_{21}} + e^{V_{22}})} \frac{e^{V_{31}}e^{V_{32}}}{\sum_{j=1}^{3} e^{V_{3j}}(e^{V_{32}} + e^{V_{33}})} \\ &\quad + \frac{e^{V_{13}}e^{V_{12}}}{\sum_{j=1}^{3} e^{V_{1j}}(e^{V_{12}} + e^{V_{11}})} \frac{e^{V_{21}}e^{V_{23}}}{\sum_{j=1}^{3} e^{V_{2j}}(e^{V_{23}} + e^{V_{22}})} \frac{e^{V_{32}}e^{V_{31}}}{\sum_{j=1}^{3} e^{V_{3j}}(e^{V_{31}} + e^{V_{33}})} \\ &= \frac{e^{V_{12}+V_{13}+V_{21}+V_{23}+V_{31}+V_{32}}}{\sum_{j=1}^{3} e^{V_{1j}} \sum_{j=1}^{3} e^{V_{2j}} \sum_{j=1}^{3} e^{V_{3j}}} \Bigg(\frac{1}{(e^{V_{11}} + e^{V_{13}})(e^{V_{21}} + e^{V_{22}})(e^{V_{32}} + e^{V_{33}})} \\ &\quad + \frac{1}{(e^{V_{11}} + e^{V_{12}})(e^{V_{22}} + e^{V_{23}})(e^{V_{31}} + e^{V_{33}})} \Bigg).\end{aligned} \tag{8}$$

The sample space $\mathcal{S} = \{A, B, C, D, E\}$ (see p. 10 for the definition of each outcome) contains all possible outcomes of the game. Now, since a merger between f_1 and f_2 is only possible in conjunction with f_3 not merging, we can write $\mathsf{P}(\mathcal{I}_{12} = 1) \equiv \mathsf{P}(\mathcal{I}_{12} =$

$\mathcal{I}_{33} = 1) = \mathsf{P}(B)$. Similar arguments lead to the following equivalences:

$$\mathsf{P}(\mathcal{I}_{13} = 1) \equiv \mathsf{P}(C)$$
$$\mathsf{P}(\mathcal{I}_{23} = 1) \equiv \mathsf{P}(D).$$

We have already derived formulas for $\mathsf{P}(B)$, $\mathsf{P}(C)$, $\mathsf{P}(D)$ (see equation (6)), and $\mathsf{P}(E)$ (see equation (8)). Therefore, using the equivalence relations above, we only need $\mathsf{P}(A)$ in order to be able to calculate the probability of each outcome in $\mathcal{S}$. But since A is the complement of $\{B \cup C \cup D \cup E\}$, we must have $\mathsf{P}(A) \equiv 1 - \mathsf{P}(B) - \mathsf{P}(C) - \mathsf{P}(D) - \mathsf{P}(E)$. This completes the treatment of the probabilities of each possible outcome, and the next section is about estimation.

4. ML estimation

For a sample of three firms, given that we observe $\boldsymbol{x} = (\boldsymbol{x}_1\ \boldsymbol{x}_2\ \boldsymbol{x}_3)$, the log-likelihood function can be written as

$$\begin{aligned}\ell(\boldsymbol{\beta}|\boldsymbol{x}) &= y_A \ln \mathsf{P}(A) + y_B \ln \mathsf{P}(B) + y_C \ln \mathsf{P}(C) + y_D \ln \mathsf{P}(D) + y_E \ln \mathsf{P}(E) \\ &= \sum_{r=1}^{5} y_r \ln \mathsf{P}_r\end{aligned}$$

where $r = A, B, C, D,$ or E, $y_r = 1$ if the outcome of the game is r and zero otherwise, and P_r were derived in the previous section. Since only one outcome is possible for a given sample of firms, the likelihood function is simply the probability of that outcome. A natural way to estimate $\boldsymbol{\beta}$ is to perform a numerical maximization of the log-likelihood function. But we have a problem, namely the few observations available. In an realistic setting, a researcher would need data on more than three firms in order to get an estimate of $\boldsymbol{\beta}$, but for say $n = 100$, $\ell(\boldsymbol{\beta}|\boldsymbol{x})$ would become extremely complicated to even write down. Thus, we need to make some additional assumptions in order to construct an estimable model.

Assume that a researcher has data on a large number of firms, but there is some additional information, namely that the firms can be divided into groups of three. The formation of groups is assumed to be exogenously given. Firms within each group are allowed to merge, but no firm is allowed to merge with a firm in another group. For instance, the total sample might consist of M groups of firms from a certain country, with each group representing a certain industry. If we are willing to assume that $\boldsymbol{\beta}$ is the same across groups, the log-likelihood for the total sample is

$$\ell_M(\boldsymbol{\beta}|\boldsymbol{X}) = \sum_{m=1}^{M} \sum_{r=1}^{5} y_{mr} \ln \mathsf{P}_{mr},$$

where $r = A, B, C, D,$ or E, $y_{mr} = 1$ if the outcome in group m is r and zero otherwise, and $\boldsymbol{X}$ contains vectors of observed data for all M groups.

There is still an issue to consider before we can put the model to work, namely how to deal with the case when no stable matching exists. In an empirical setting, it might be the case that distinguishing between the events A and E is impossible, simply because the non-existence of a stable matching might be impossible to find in a database. Generally, we do not expect everything to crash if no stable matching exists; instead, firms simply keep on operating on their own. On the other hand, one could in principle think of data where A and E are distinguishable. For instance, data on a seemingly never-ending round of bid proposals among a group of firms could be a sign of instability.

Depending on whether one can or cannot distinguish between A and E, the log-likelihood function takes on different forms. If there is a possibility to distinguish between the two events, the log-likelihood function defined above can be used. But in the more realistic case when it is impossible to detect instability, the set of possible outcomes has to be redefined as $\mathcal{S} = \{A \cup E, B, C, D\}$. Rewriting the log-likelihood function of a particular firm group to be in accordance with this results in:

$$\begin{aligned}
\ell(\boldsymbol{\beta}|\boldsymbol{x}) &= y_{AE} \ln \mathsf{P}(A \cup E) + y_B \ln \mathsf{P}(B) + y_C \ln \mathsf{P}(C) + y_D \ln \mathsf{P}(D) \\
&= y_{AE} \ln \{1 - [\mathsf{P}(B) + \mathsf{P}(C) + \mathsf{P}(D) + \mathsf{P}(E)] + \mathsf{P}(E)\} \\
&\quad + y_B \ln \mathsf{P}(B) + y_C \ln \mathsf{P}(C) + y_D \ln \mathsf{P}(D) \\
&= y_{AE} \ln \{1 - \mathsf{P}(B) - \mathsf{P}(C) - \mathsf{P}(D)\} + y_B \ln \mathsf{P}(B) + y_C \ln \mathsf{P}(C) + y_D \ln \mathsf{P}(D),
\end{aligned}$$

where $y_{AE} = 1$ if the outcome is either A or E, and zero otherwise. Accordingly, the log-likelihood of the whole sample of M groups can be written as

$$\ell_M(\boldsymbol{\beta}|\boldsymbol{X}) = \sum_{m=1}^{M} \sum_{r=1}^{4} y_{mr} \ln \mathsf{P}_{mr}, \tag{9}$$

where $r = AE, B, C,$ or D, $y_{mr} = 1$ if the outcome in group m is r and zero otherwise, and $\boldsymbol{X}$ contains vectors of observed data for all M groups.

Let $\widehat{\boldsymbol{\beta}}(\boldsymbol{X}) \equiv \widehat{\boldsymbol{\beta}}$ be the maximum likelihood estimator (MLE) of $\boldsymbol{\beta}$, i.e., the parameter value at which $\ell_M(\boldsymbol{\beta}|\boldsymbol{X})$ attains its maximum as a function of $\boldsymbol{\beta}$, with $\boldsymbol{X}$ held fixed. Being a MLE, well-known results state that $\widehat{\beta}$ is consistent and asymptotically efficient.[4] Asymptotic efficiency implies that we can approximate the true asymptotic variance of the MLE with the Fisher information matrix. Below we will come back to this issue with an estimation of the error magnitude when using the asymptotic variance in finite samples. The regularity conditions that each $\ell(\beta|\boldsymbol{x})$ has to fulfill in order for the MLE

[4] See Casella and Berger (2002, Theorems 10.1.6 and 10.1.12).

to be consistent and asymptotically efficient are nicely summarized in Casella and Berger (2002, section 10.6.2).

In a realistic setting, samples are always finite, and might in addition be quite small. Below is a simulation study of the performance of the MLE in such a case. The general idea is to generate data according to the firm-matching game described above, and perform a numerical maximization of the log-likelihood function, thus obtaining $\widehat{\beta}$. Doing this a large number of times provides us with means of analyzing the sampling properties of the MLE.

As above, M is the number of groups, and there are three firms in each group. The first step in each replication is to generate $\Pi_{ij} = \beta x_{ij} + \varepsilon_{ij}$ for each firm group, where β is set to $0.1, 0.25, 0.5$, or 1, $x_{ij} \sim i.i.d.\ N(0,1)$, $\varepsilon_{ij} \sim$ extreme value with location parameter equal to zero and scale parameter equal to one, and $i, j = 1, 2, 3$. Firm attributes are the same for each group and replication, while the stochastic parts of the indices vary. When the indices are generated, the firms play the matching game and the Roth-Vande Vate algorithm leads to a matching. Finally, the log-likelihood function in (9) is maximized, resulting in $\widehat{\beta}$.

The above is repeated $R = 50,000$ times. Index the parameter estimate from each replication as $\widehat{\beta}_r$, the corresponding vector with attributes as $\boldsymbol{X}_r$, and define $\widehat{H}_r$ to be the Hessian matrix evaluated at $\widehat{\beta}_r$, where $r = 1, 2, \ldots, R$. The finite-sample properties of the MLE will be examined using the results from the simulation with the help of the following statistics:

$$
\begin{aligned}
\overline{\beta} &= R^{-1} \sum_{r=1}^{R} \widehat{\beta}_r \\
\widehat{\beta}_{max} &= \max(\widehat{\beta}_1, \widehat{\beta}_2, \ldots, \widehat{\beta}_R) \\
\widehat{\beta}_{min} &= \min(\widehat{\beta}_1, \widehat{\beta}_2, \ldots, \widehat{\beta}_R) \\
\widehat{\sigma}^2_{\widehat{\beta}} &= R^{-1} \sum_{r=1}^{R} (\widehat{\beta}_r - \overline{\beta})^2 \\
\widehat{\mathcal{I}}(\beta|\boldsymbol{X}) &= R^{-1} \sum_{r=1}^{R} (-M^{-1} \widehat{H}_r^{-1}) \equiv -R^{-1} M^{-1} \sum_{r=1}^{R} \left\{ \left. \frac{\partial \ell(\widehat{\beta|\boldsymbol{X}_r})}{\partial \beta^2} \right|_{\widehat{\beta}_r} \right\}^{-1}
\end{aligned}
$$

The mean of the estimators in the simulation, $\overline{\beta}$, should be close to β for an unbiased estimator. The MLE's bias should decrease with the sample size and asymptotically go to zero. $\widehat{\mathcal{I}}(\beta|\boldsymbol{X})$ is the mean estimate of the Fisher information, i.e., a lower bound of the variance of the best unbiased estimator of β.

		M			
β		25	50	100	1,000
0.1	$\overline{\beta} - \beta$	0.0212	0.008	-0.0013	0.0000
	$\widehat{\beta}_{max}$	1.5593	1.1392	0.5725	0.1981
	$\widehat{\beta}_{min}$	-0.9557	-0.9212	-0.5330	0.0018
	$\widehat{\sigma}^2_{\widehat{\beta}}$	0.060170	0.046883	0.013566	0.000565
	$\widehat{\mathcal{I}}(\beta\|\boldsymbol{X})$	0.055909	0.045834	0.013380	0.000558
0.25	$\overline{\beta} - \beta$	0.0265	0.0288	0.0049	0.0003
	$\widehat{\beta}_{max}$	1.6849	2.4537	0.8707	0.4919
	$\widehat{\beta}_{min}$	-0.8784	-0.6833	-0.1525	0.0023
	$\widehat{\sigma}^2_{\widehat{\beta}}$	0.090244	0.064366	0.011374	0.002835
	$\widehat{\mathcal{I}}(\beta\|\boldsymbol{X})$	0.083374	0.060727	0.011301	0.002812
0.5	$\overline{\beta} - \beta$	0.0107	0.006	0.0095	0.0001
	$\widehat{\beta}_{max}$	2.1167	1.8049	1.5515	0.6453
	$\widehat{\beta}_{min}$	-0.7000	-0.5660	-0.2590	0.3635
	$\widehat{\sigma}^2_{\widehat{\beta}}$	0.080059	0.066275	0.035656	0.001082
	$\widehat{\mathcal{I}}(\beta\|\boldsymbol{X})$	0.077834	0.065906	0.035011	0.001080
1	$\overline{\beta} - \beta$	0.0326	0.0229	0.0065	0.0015
	$\widehat{\beta}_{max}$	4.9656	3.2199	1.5381	1.3205
	$\widehat{\beta}_{min}$	-1.8604	-0.5486	0.6496	0.7329
	$\widehat{\sigma}^2_{\widehat{\beta}}$	0.394940	0.134740	0.012067	0.004114
	$\widehat{\mathcal{I}}(\beta\|\boldsymbol{X})$	0.253540	0.130160	0.011799	0.004108

Table 1: Sampling properties of the MLE for various values of the true parameter β and of the number of groups M; 50,000 replications.

The results from the simulation were generated using Ox version 3.40 and the BFGS algorithm was used when maximizing the log-likelihood functions.[5] The initial parameter guess was set to zero, but initial values ranging from -5 to 5 were tested and not found to alter the results.

To illustrate, Figure 1 shows the log-likelihood function for data generated with $\beta = 0.5$, and for $M = 100$ and 1,000, respectively. In each graph, the function is evaluated at 1,000 different parameter values. The left-hand graphs show the function value on the parameter range $[\widehat{\beta} - 5, \widehat{\beta} + 6]$, while the ones on right-hand show the function on $[\widehat{\beta} - 0.1, \widehat{\beta} + 0.1]$. This graphical examination of the log-likelihood function gives an indication of its concavity, which in turn explains the apparent insensitivity of the maximization algorithm to different starting values.

Now, consider the results presented in Table 1. The average of the estimators, $\overline{\beta}$, behaves as predicted over all values of β, namely converging to the true parameter value with M increasing. Even for sample sizes of 25 we can expect the MLE to be acceptably close to β, and when $M = 1,000$, the difference is negligible.

[5]See for instance Judd (1998, p. 114) for a description of the algorithm.

Also the span between the extrema of the simulated distribution of $\widehat{\beta}$ decreases with M. But even for $M = 1,000$, a particular estimator can differ from the true value with between as much 100 % (for $\beta = 0.1$) and 30 % (for $\beta = 1$). However, one must keep in mind that these numbers say nothing about the probability of getting a parameter estimate that far from the true value. To be able to see the whole picture, we need to look at the whole distribution of the estimates. To exemplify, in Figure 2 we show the simulated density of $\widehat{\beta}$ with $\beta = 0.5$ and $M = 1,000$. The bars represent the normal distribution with the same mean and variance, shown as a comparison. First, note that extreme values like $\widehat{\beta}_{max} = 0.65$ and $\widehat{\beta}_{min} = 0.36$ are very unlikely. Second, a visual comparison with the density of the normal distribution reveals a great resemblance, which asymptotically is to be expected of a MLE.

Next, consider the variance of the MLE. The estimated Fisher information can be used as a measure of a lower bound of the variance of $\widehat{\beta}$, and from theory, we know that it is asymptotically achieved by the MLE. As we can see, for $M = 1,000$ the lower bound is almost reached. Thus, for practical purposes and for large M, the numerically calculated Fisher information can be used as an estimate of the true variance of $\widehat{\beta}$. Note, however, that this underestimates the true variance, and the underestimation is not negligible for $M \leq 100$. In hypothesis testing, using the asymptotic variance estimate for small samples would imply a tendency to reject the null too often.

To sum up, our simulation suggests that the MLE proposed in the previous section has quite good finite-sample properties. For large samples, Fisher information can well be used as a variance estimate. Furthermore, the percentage bias is small for $M \geq 100$, and acceptable for smaller samples.

5. Concluding remarks

This paper proposes a theoretical model of mergers based on individual firm behaviour. The model can be used directly in applied work, and using simulations, it is shown that in a laboratory environment, the proposed maximum-likelihood estimator has several desirable properties in finite samples, and is thus readily applicable to empirical data.

The way it is presented the model has several built-in restrictions, the most obvious being the assumption of solely three firms in each group. Although a generalization to larger groups is straightforward, the resulting expressions for different outcomes are bound to be complex and time-demanding to derive. Thus, an obvious path for further research is to find more general formulas for the probability of different outcomes with more than three firms in each group.

A different approach consists of evading the problem by changing the rules of the roommate game, such that the likelihood function becomes less complicated. This in practice means that we abandon the roommate game, but in return get a model more suitable for empirical estimation. Depending on the new game rules, the merger model might still be economically sensible.

Still one direction for future research is to evaluate the properties of the MLE analytically. A formal check of consistence and asymptotic efficiency (i.e., a check of the regularity conditions on each $\ell(\boldsymbol{\beta}|\boldsymbol{x})$, stated in Casella and Berger (2002, Section 10.6.2)), and a check for global concavity of the log-likelihood function might be useful.

References

Casella, G., and R. L. Berger (2002): *Statistical Inference.* Duxbury Press, Pacific Grove, CA, USA.

Chung, K.-S. (2000): "On the existence of Stable Roommate Matchings," *Games and Economic Behavior*, 33, 206–230.

Diamantoudi, E., E. Miyagawa, and L. Xue (2004): "Random Paths to Stability in the Roommate Problem," *Games and Economic Behavior*, 48, 18–28.

Ericson, R., and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62, 53–82.

Fridolfsson, S.-O., and J. Stennek (2004): "Industry Concentration and Welfare – On the Use of Stock Market Evidence from Horizontal Mergers," Working Paper from The Research Institute of Industrial Economics, IUI.

——— (2005): "Why Mergers Reduce Profits and Raise Share Prices – A Theory of Preemptive Mergers," *Journal of the European Economic Association*, 3, 1083–1104.

Gale, D., and L. Shapley (1962): "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 69, 9–15.

Gowrisankaran, G. (1999): "A Dynamic Model of Endogenous Horizontal Mergers," *The RAND Journal of Economics*, 30, 56–83.

Griliches, Z., and M. D. Intriligator (eds.) (1984): *Handbook of Econometrics*, vol. 2. Elsevier, Amsterdam.

Horn, H., and L. Persson (2001): "Endogenous Mergers in Concentrated Markets," *International Journal of Industrial Organization*, 19, 1213–1244.

Irving, R. W. (1985): "An Efficient Algorithm for the 'Stable Roommates' problem," *Journal of Algorithms*, 6, 577–595.

Judd, K. L. (1998): *Numerical Methods in Economics.* MIT Press, Cambridge, Massachusetts.

Kamien, M. I., and I. Zang (1990): "The Limits of Monopolization Through Acquisition," *The Quarterly Journal of Economics*, 105, 465–499.

McFadden, D. (1974): *Conditional Logit Analysis of Qualitative Choice Behavior* chap. 4, pp. 105–142, in (Zarembka 1974).

——— (1984): *Econometric Analysis of Qualitative Response Models* chap. 24, pp. 1396–1457, vol. 2 of (Griliches and Intriligator 1984).

Roth, A. E., and M. A. O. Sotomayor (1990): *Two Sided Mathching: A Study in Game-Theoretic Modeling and Analysis.* Cambridge University Press, Cambridge.

Roth, A. E., and J. H. Vande Vate (1990): "Random Paths to Stability in Two-Sided Matching," *Econometrica*, 58, 1475–1480.

Tan, J. J. M. (1991): "A Necessary and Sufficient Condition for the Existence of a Complete Stable Matching," *Journal of Algorithms*, 12, 154–178.

Zarembka, P. (ed.) (1974): *Frontiers in Econometrics.* Academic Press, New York.

Appendix A. Figures

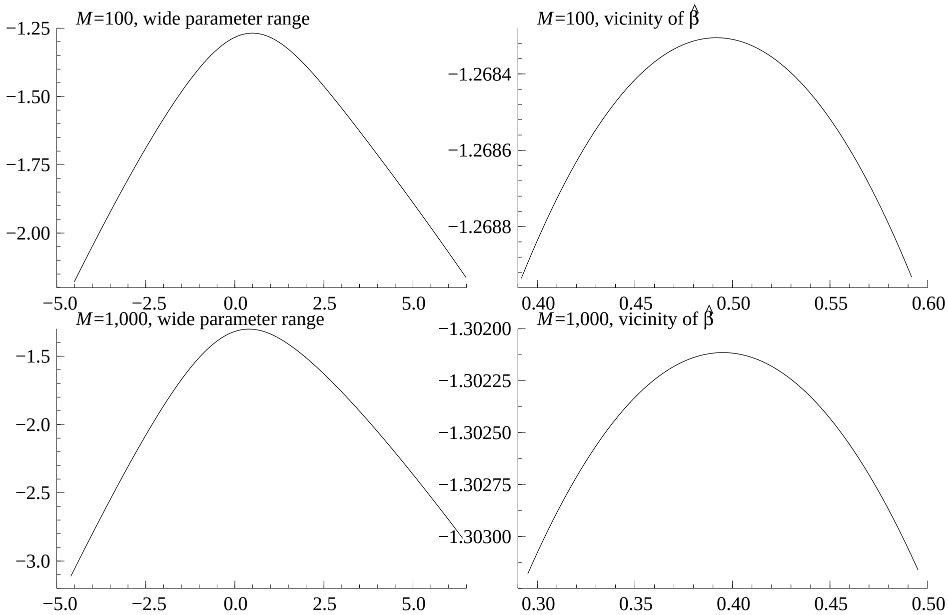

Figure 1: The log-likelihood function at 1,000 parameter values in each graph, $\beta = 0.5$.

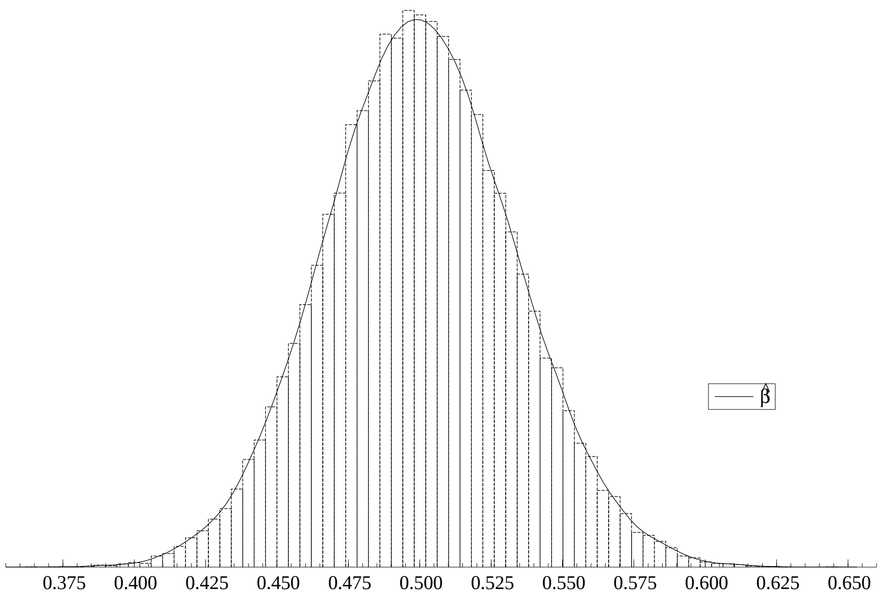

Figure 2: Simulated density of $\widehat{\beta}$ for $\beta = 0.5$ and $M = 1,000$, using 50,000 replications. Density of the normal distribution with the same mean and variance as a reference.