Transportation Research Part B 108 (2018) 249–260

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Understanding many-to-many matching relationship and its correlation with joint response

Dapeng Zhang[a], Xiaokun (Cara) Wang[b,*]

[a] Department of Civil and Environmental Engineering, 5304 JEC Building, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590 USA

[b] Department of Civil and Environmental Engineering, 4032 JEC Building, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590 USA

ARTICLE INFO

Article history:
Received 20 February 2017
Revised 18 December 2017
Accepted 20 December 2017

Keywords:
Many-to-many
Joint response
Matching model
Sample selection
Bayesian MCMC

ABSTRACT

Many-to-many matching relationship in a two-sided market has been widely observed in today's transportation activities. Observation of such matching relationship raises some interesting questions: what factors drive the matching of two agents? Is the formation of matching relationship related with joint behavior which may lead to different understandings of planning and operation? To answer these questions, econometric models may be the best methodology. However, to the authors' best knowledge, there lacks a well-established econometric model to explain the observed data that contains matching relationship in a two-sided transportation market. Therefore, this paper proposes an innovative ordinal joint response model to bridge the gap. The proposed model consists of two regression equations: the first uses a latent dependent variable to disentangle the many-to-many matching relationship; the second specifies an ordered probit equation to investigate the ordinal outcome of joint behavior. Error terms of the two equations are assumed correlated to capture the correlation of the matching process and joint behavior. An example of airline-airport matching is used to demonstrate the proposed model.

1. Introduction

Many-to-many matching relationship in a two-sided market has been widely observed in today's transportation activities. For example, in a supply chain segment, upstream agents match with downstream agents to deliver cargo and each agent is working with multiple counterparties at the same time. In an aviation industry, airports host several airlines and airlines use many airports for their daily operation. A formal definition of a many-to-many matching is that more than one agents on one side of a market match with more than one agents on the other side of the market. Observation of such matching relationship raises some interesting questions: what factors drive the matching of two agents? Is the formation of matching relationship related with joint behavior of matched agents?

Econometric models, which can capture the correlation between a set of attributes and outcomes, seem to be the best method to answer these questions. Throughout the transportation literature of econometric modeling, models that capture matching relationship have not been well investigated. Other fields, such as finance, have adopted two-sided matching models to explain the bank-firm matching relationship, although the literature is scarce. An example is Sorensen (2007), which

* Corresponding author.
E-mail addresses: zhangdapeng@live.com (D. Zhang), wangx18@rpi.edu (X. (Cara) Wang).

https://doi.org/10.1016/j.trb.2017.12.011
0191-2615/

uses one matching equation to explain the bank-firm partner selection and one binary outcome equation to formulate firms' IPO choices. The matching structure is one-to-many so that one company can be invested by multiple banks but one bank can invest in only one company. The developed framework assumes that a firm starts with assessing the characteristics of all banks that can potentially provide the IPO service and, based on the assessment, selects the banks that fulfill the company's need. If the company's most favorable banks dislike the company, the company turns to the next best bank until the desired bank also likes the company. As the partner selection process is bidirectional, a similar assessment process occurs on the bank's side at the same time. Sorensen's model (2007) can identify coefficients of influential variables that yield the observed matching relationship, which enlightens the development of this research that borrows Sorensen's idea and improves the model specification to solve transportation problems.

This paper uses a case study of flight delay problems to illustrate the proposed model. Airlines and airports are assumed to be the agents in the two-sided market. An airline and an airport are paired as long as one passenger flight of the airline company lands at the airport. The first equation of the proposed model can identify the influential factors that yield their many-to-may matching relationship. The second equation quantifies flight delay as an ordinal outcome and investigate the influential factors of flight delay. The correlation of the two equations is specified by a more flexible variance-covariance matrix. The proposed model is estimated by a Bayesian Markov Chain Monte Carlo (MCMC) approach, tested by a set of simulations, and applied to analyze flight on-time performance with collected data.

This paper will be organized as follows. The next section will review literature in the matching model and airline-airport collaboration. Then, the proposed model will be introduced by its specification and estimation. A set of simulation studies will be followed to show the model's performance. Finally, a flight on-time performance dataset will be analyzed with the proposed model and followed by conclusions.

2. Literature review

2.1. Matching model

Methodologies analyzing matching relationship have a relatively short history. One of the most classic matching problems is the marriage problem where decision makers look for the best counterparty in a two-sided market. A classic marriage problem aims at finding a stable matching between two equally-sized sets of agents given a ranking of preferences for each agent. A matching is "stable" when a man and a woman are both engaged, but not to each other, and upon the completion of the matching, it is not possible for them to prefer each other over their current partners. In 1962, Gale and Shapley (1962) tried to find whether there is a stable way to match men and women so that no unmatched pairs are left. The Gale-Shapley algorithm proved that it is always possible to solve the stable matching problem and make all marriages stable for any equal number of men and women. Gale and Shapley's work was a fundamental study that started the 60 years' matching market studies in finding solutions to real-world matching problems. Their work was recently honored by the Nobel Prize in Economic Sciences in 2012 (The Nobel Foundation, 2012). Some of the remarkable work regarding matching among agents including Roth and Sotomayor (1992). However, none of them looked at the problem from the perspective of econometrics, which investigates the effects of influential factors on forming the observed matching relationship. Observed matching data include which firms do business with which firms, which men are married with which women, and which players are teammates with which players, among other contexts. The basic economic idea is that one individual would like to partner with the most attractive partners, leading to highest benefits for the individual. Econometricians seek to understand influential factors that determine the observed matching and estimate the parameters of these factors.

An important literature to this paper is Sorensen (2007), which uses a two-sided matching model to the bank-firm matching and its impact on firms' initial public offering. It specifies matching utility of all possible pairs in the matching process and analyze joint response for only matched pairs. A Bayesian MCMC approach is employed to estimate parameters of factors in determining all parameters in the two equations. This framework evaluates each pair's matching utility to analyze the selection of partners. Such a method is a fundamental work for this paper. Similar econometric matching models have been discussed in a limited number of empirical studies. Chen (2013) specifies the utility equations for each side of the partner to analyze the premium of bank loans. This study is an extension of Sorensen (2007) in which the utility of paired partners is investigated instead of utility of decision makers, respectively. However, this model may suffer from identification problems if extended to discrete outcomes. To the author's best knowledge, no other studies have conducted the research in a similar way. Other notable studies have analyzed the matching process from the market's perspectives, using different estimation methods, or without considering the mutual selection process. For example, Choo and Siow (2006) investigate the stable matching relationship from the market's perspective rather than the each decision maker's perspective. Similar marriage analyses can be also found in Siow (2008). Hitsch et al. (2010) also analyze a marriage dataset, but the methodology does not consider the mutual preference by sorting pairwise utility. Fox (2008) and Levine (2009) use maximum score estimators (e.g., a non-parametric model) to identify parameters in the matching model.

Sorensen's model (2007) consists of two equations with one matching equation explaining the observed matching structure and the other binary outcome equation modeling joint responses. The two equations are connected by a variance-covariance matrix that captures the simultaneity of the two data generating processes. The matching equation is identified by a set of inequality conditions of pairwise utility that is developed based on the preference of matching. Such a model structure will be directly borrowed to this transportation matching research, because the assumptions of bank-firm rela-