Who Gets Placed Where and Why? An Empirical Framework for Foster Care Placement

Alejandro Robinson-Cortés*

February 25, 2021

Abstract

This paper presents an empirical framework to study the assignment of children into foster homes and its implications on placement outcomes. The empirical application uses a novel dataset of confidential foster care records from Los Angeles County, CA. The estimates of the empirical model are used to examine policy interventions aimed at improving placement outcomes. In general, it is observed that market thickness tends to improve expected placement outcomes. If placements were assigned across all the administrative regions of the county, the model predicts that (i) the average number of foster homes children go through before exiting foster care would decrease by 8% and (ii) the distance between foster homes and children's schools would be reduced by 54%.

*Department of Economics, University of Exeter Business School (e-mail: a.robinson-cortes@exeter.ac.uk). I am extremely grateful to my advisors Matt Shum, Leeat Yariv, and Federico Echenique, for all their guidance, patience, and constant support. I thank Ben Gillen, Bob Sherman, Jean-Laurent Rosenthal, Laura Doval, Lucas Núñez, Marcelo Fernández, Mike Ewens, Sergio Montero, Wade Hann-Caruthers, Yimeng Li, and participants of the North American Summer Meetings of the Econometric Society at UC Davis 2018, the USC Children Data Network Brown Bag Seminar, and Caltech's Applied Micro Lunch and Graduate Proseminar for helpful discussions and suggestions. I'm also grateful to Maricruz Arteaga-Garavito for telling me about foster care and sparking my interest. The author obtained IRB approval and a limited waiver of confidentiality from the Juvenile Division of the Superior Court of California to analyze confidential records of the Department of Children and Family Services of Los Angeles County. The analyses and interpretations of all the data used in this study are the sole responsibility of its author. The aforementioned institutions and their agents or employees bear no responsibility for the analyses and interpretations presented here.

1 Introduction

The assignment of scarce resources is at the heart of economics. In this paper, I study one particular assignment setting that has been largely overlooked by the economics literature—the placement of children into foster homes. I develop an empirical framework that captures how social workers match children and foster homes in the field. The analysis centers on the relationship between placement assignments and outcomes.

I estimate an econometric model using a novel dataset of confidential county records at the micro-level from the largest foster care system in the United States, the one in Los Angeles County, California. Motivated by the literature on children welfare studies (and anecdotal evidence from conversations with social workers), my definition of *placement outcomes* includes both the duration of placements and whether they are disrupted (i.e., children are moved from one foster home to another) or terminate because children exit foster care.

I use the estimates of the model to examine various policy interventions aimed at improving placement outcomes. I find that thicker markets generate better outcomes in the sense that they result in lower disruption rates, but the effects are different along different dimensions. Specifically, the model predicts that the gains from assigning placements across geographic regions in the county are greater than those generated by delaying assignments. Counterfactual exercises show that pooling the assignments across all the regional offices in the county would decrease the expected number of placements each child goes through before exiting foster care by 8%. I also quantify the system-wide effects of specific types of foster homes. I find that increasing the share of placements involving children's relatives (also known as kinship care) would lead to lower placement disruption rates and longer placements. In contrast, the model predicts mixed effects from increasing the share of foster homes that are recruited and trained by non-profit agencies.

The model is designed to capture the co-dependence between placement assignments and outcomes. On the one hand, the model captures how the assignments of placements are driven by their expected outcomes. On the other hand, it also recognizes that the outcomes observed in the data are selected through such assignment. The interplay between assignments and outcomes causes an endogeneity problem. The matching mechanism determines which placement outcomes are observable. Hence, the observed distribution of placement outcomes is biased inasmuch as placement assignments are driven by unobservables correlated with outcomes. To identify the true distribution of outcomes, the model exploits the exogenous variation across the dates and geographic regions in which children enter foster care. I study matching markets at the daily level across the nineteen administrative regions defined by the Los Angeles County Department of Children and Family Services.

It is widely recognized that stable foster care placements are essential for the social, emotional, and cognitive development of children (UC Davis 2008). Social workers in the field also strive to assign long-lasting placements to minimize future workloads. Nonetheless, it fairly common that children go through multiple foster homes while they are in foster care.[1] A key factor affecting outcomes is the placement characteristics, which are the result of the assignment of children into foster homes. Understanding how children are being assigned to foster homes allows one to analyze how the matching mechanism used in the field translates into outcomes via placement characteristics. For example, the estimates of the model show that the gains from thicker markets come largely from being able to assign children to foster homes that are closer to their schools. The model predicts that if the assignments of placements were determined at the county-level (instead of within geographic regions), the average distance between children's schools and their foster homes would be cut by 54%.

I model the assignment of children into foster homes as an optimal matching problem, and I model placement outcomes with a mixed competing risks duration model. The matching problem allows for idiosyncratic variation in the preferences of children over foster home characteristics, and vice versa. At the same time, it takes into account that placements are assigned on the basis of their expected outcomes. I model unobservable heterogeneity through frailty terms in the outcome distribution. To account for possible selection bias (i.e., that placements may be assigned because of unobservables correlated with outcomes), I assume that the decision-maker choosing the matching between children and foster homes observes such

[1] For example, of all the children who exited foster care in the U.S. during 2015, 56.1% of them went through at least two placements, and the average number of placements per child was 2.56 (NDACAN 2015). It has also been shown that the time children spend in foster care, as well as the number of placement disruptions they experience, are associated in adult life with emotional and behavioral difficulties, increased criminal convictions, and higher depression and smoking rates (Dregan and Gulliford 2012).

frailty terms. Thus, the distribution of outcomes generated by the model is conditional on the assignment chosen and incorporates unobservable heterogeneity.

The estimates of the matching model allow me to quantify the trade-offs that social workers incur when assigning placements. For instance, at first sight, it seems intuitive that social workers aim to assign the placements that are expected to have the longest durations in order to avoid placement disruptions. However, this reasoning ignores the intimate co-dependence between a placement's duration and its termination reason. Indeed, according to the model estimates, social workers' assignments reflect a dislike for duration conditional on termination reason. That is, if a placement were known to be disrupted, the model estimates indicate that social workers would prefer it to be disrupted sooner rather than later. Similarly, if it were known that a placement will terminate because the child will exit foster care to a permanent placement, social workers would prefer this to happen as soon as possible. At the same time, the estimates show that social workers prioritize minimizing disruptions over placement duration. That is, regardless of a placement's duration, the model predicts that social workers would always prefer for placements *not* to be disrupted.

The paper is organized as follows. I review the related literature in what remains of the introduction. In Section 2, I provide an institutional background of foster care, and describe the data. Section 3 presents the econometric model. In Sections 4 and 5, I discuss the identification of the model and the estimation technique. Section 6 reports the estimation results. Section 7 shows the results of the counterfactual exercises, and Section 8 concludes.

Related Literature.— The main contribution of this paper is to develop an empirical framework to study (1) how children are assigned into foster homes, and (2) how the matching mechanism underlying such assignment translates into placement outcomes. Slaugh, Akan, Kesten, and Ünver (2015) is the only other paper in the literature that applies tools from matching and market design to a question related to foster care. They analyze the Pennsylvania Adoption Exchange program, whose main aim is to facilitate the adoption of foster children through a computerized recommendation system. They analyze the effect that improvements to the system—in terms of enhancing the capacity of social workers to match children and prospective adoptive parents—have on the rate of successful adoptions.

Baccara, Collard-Wexler, Felli, and Yariv (2014) analyze data from an online platform that seeks to facilitate adoptions. Although they are distinct in fundamental ways, adoption and foster care closely related. Parents who are seeking to adopt often become foster parents beforehand, and, in many cases, foster children are adopted by their foster parents. Baccara, Collard-Wexler, Felli, and Yariv (2014) focus on the preferences that prospective adoptive parents show for children. They find a favorable preference for girls, and a preference against African Americans.

Overall, the economics literature analyzing questions related to foster care is slim. In a series of papers, Doyle Jr. 2007; 2008; 2013 aims to evaluate the impact of foster care on long-term outcomes. Their approach exploits that, in many cases, social workers are assigned randomly to investigate reports of abuse and neglect. This random assignment allows them to identify the "treatment effect" of foster care on schooling, employment, and criminality. Doyle Jr. and Peters (2007) use variation in the subsidies offered to foster parents to estimate the supply curve of foster homes. Analyzing data from the late 1980s early 1990s, they estimate that, in states with shortages of foster homes, an increase in subsidies by 10% increases the quantity supplied by 3%.[2]

In broader terms, this paper belongs to the empirical matching and market design literature (Roth 2016). The common denominator in this literature is the formulation and estimation of structural models that incorporate key institutional aspects of the market being studied. In a seminal contribution, Choo and Siow (2006) study the marriage market in a transferable utility (TU) environment. Their setup is based on the Assignment Game developed by Shapley and Shubik (1971). See Graham 2011; 2013, Chiappori, Oreffice, and Quintana-Domeque (2012), and Galichon and Salanié (2015) for extensions and generalizations of their approach. Choo (2015) further extends the analysis to a dynamic setting. More generally, Fox (2016) studies nonparametric identification and estimation of TU matching markets. Buchholz (2019) and Fréchette, Lizzeri, and Salz (2019) study matching models in the market for taxis.[3]

In a non-TU environment, Agarwal (2015) formulates and estimates a matching model of the medical match (NRMP). Agarwal and Somaini (2018) study the

[2]See Doyle Jr. and Aizer (2018) for an excellent literature review on the current state of empirical work in economics on child maltreatment and its relation to foster care and intimate partner violence.

[3]Market-clearing transfers need not only be monetary prices (e.g. passengers waiting for taxis "pay" in waiting-time units), see Galichon and Hsieh (2017).

strategic incentives of different mechanisms in the assignment of children to public schools. For other recent contributions to the empirical study of school choice, see Narita (2016), Hwang (2016), Calsamiglia, Fu, and Güell (2017), and Abdulkadiroğlu, Agarwal, and Pathak (2017). There is also a growing literature analyzing kidney exchange (e.g. Agarwal et al. 2017), waiting-list mechanisms for organ donation (Agarwal et al. 2019) and public housing allocation (Waldinger 2019).

All the studies cited in the previous two paragraphs model assignments according to specific matching mechanisms. The TU literature generally assumes that the market is cleared via equilibrium transfers. In non-TU environments, the assignment usually results from predetermined matching algorithms.[4] The main differences from previous studies and this paper is that the assignment mechanism underlying foster care neither involves equilibrium transfers nor makes use of a systematic matching algorithm. The matching between children and foster homes is centralized and the consequence of both (i) specific regulations and (ii) discretionary choices made by social workers in the field.

The insights from this paper may also be relevant for the growing literature on dynamic matching. One of the main objectives of this literature is to study the dynamic trade-offs between waiting time, thickness, incentives, and match quality. For notable examples, see Baccara, Lee, and Yariv (Forthcoming), Ünver (2010), Akbarpour, Li, and Gharan (Forthcoming), Doval (2018), and Ashlagi, Jaillet, and Manshadi (2013). Specifically, this paper provides an example in which increasing market thickness by delaying placements does not have sizable effect on outcomes.

2 Institutional Background and Data

2.1 Foster care in the U.S. and Los Angeles County

Every year more than a half million children go through foster care in the United States. Foster children are a particularly vulnerable population: most of them are

[4] The study of matching algorithms dates back to Gale and Shapley (1962), who formulated the well-known Deferred Acceptance (DA) algorithm. Roth (1984) documents the history of the medical match and, more specifically, how it came to employ the DA algorithm before the findings of Gale and Shapley. Given the attractive features of DA (stability and strategy-proofness), it has been proposed as a mechanism to match children to schools (Abdulkadiroğlu and Sönmez 2003). A significant portion of the school choice literature compares the DA algorithm with the so-called Boston algorithm (e.g. Abdulkadiroğlu, Pathak, Roth, and Sönmez 2005).

in foster care because they were abused, neglected, or abandoned (NDACAN 2015). The main goal of foster care is to provide temporary care for children until permanent placements can be arranged for them. When a child is moved from a foster home to a permanent placement, it is said that she exits foster care to "permanency". Children who exit to permanency usually go back to live with their birth families, or, if this is not possible, are adopted or assigned guardians. When permanent placements cannot be arranged for children, they stay in foster care until they become of age, and are emancipated from the system (also known as "aging out").[5]

The administrative management of the foster care system is at the county level in the U.S.[6] The child protection agency of Los Angeles County is the Department of Children and Family Services (DCFS).[7] As other child protection agencies, DCFS is responsible of processing and investigating reports of child abuse, taking cases to court, and implementing court resolutions. After receiving a report, county social workers conduct an investigation to determine if children need to be removed from home. The decision whether a child should be removed or not needs to be approved by a judge. The procedures regarding the investigation and removal decision are independent from placement assignment procedures. Foster care placements are assigned and managed within nineteen regions across the city of Los Angeles. When a child enters foster care, its case is handled by the regional office corresponding to the region where the child's birth mother lives. Social workers from that regional office are responsible for finding a suitable placement for the child, and overseeing her case while she remains in foster care.

2.2 Placement Assignment in Foster Care

By law, there a few factors that social workers must consider when assigning placements: (1) whether a child has relatives who are available to take care of them, in

[5]Foster care is inherently different from adoption. In general, adoptive parents have the same rights and obligations over their children as biological parents. By contrast, foster parents have very limited say in the placement of foster children. Whether a child is removed from home, placed in or exits foster care, is a decision made by the courts, who rely heavily on the input of social workers.

[6]In some cases, there is a single child protection agency for all the counties covering the same urban area (e.g., in New York City there is a single agency for the five boroughs).

[7]Specific foster care regulations vary at the state and county level. In California, the main regulations of the foster care system are provided in the Welfare and Institutions Code (WIC 2019), and the Family Code (FAM 2019). In Los Angeles County, foster care regulations are provided in the Child Welfare Policy Manual of DCFS (2019). For a history of the foster care system in the United States, see Rymph 2017.

which case children must be placed with their relatives; (2) the location of the foster home: social workers must make efforts to place children in foster homes that are near their schools and their family homes (from where they were removed), and (3) whether a child has siblings who are also in foster care, in which case efforts should be made to place siblings together.[8] However, the law does not provide a systematic way in which these factors are to be waged against one another. The law also gives social workers the discretion to assign placements that bypass these guidelines if they consider that such placements are not in the child's best interest. Likewise, children who are 10 years or older also have the right to make a brief statement in court regarding the placement decision.

In the field, social workers aim to find placements that fulfill all the requirements stated in the law, and are also suitable for children in more practical ways. For example, when evaluating prospective foster homes, they may take into account scheduling and transportations considerations, the family environment of the foster home (e.g., the age and gender of the family's biological children), and other idiosyncratic factors such as the experience of the foster parents and the history of a child in the system. The reason for taking into account each of these factors is because a main concern of social workers is for placements to be disrupted. Placements are usually disrupted because the foster family and the foster child are not able to establish a harmonious and stable relationship (e.g., the child presents behavioral problems the family is not prepared to deal with, the situation of the family changes, or problems emerge between the foster child and the family's biological children). When placements are disrupted, children need to be moved to new foster homes. In LA County, on average, foster children go through 2.1 foster homes before exiting to permanency.

I gathered the above observations through informal conversations with a handful of social workers with experience in the field. Overall, my impression from these conversations is that apart from the guidelines embedded in the law, social workers work on a case-by-case basis. They treat each case differently, and wage all of the involved factors in a case to find the best possible placement. Another common observation is that, in many cases, ideal placements are just not possible because of the shortage of foster homes. As children enter foster care, social workers within each

[8] See DCFS 2019, Sec. 0100-510.60; FAM 2019, Div. 12, Part 6, Sec. 7950, and WIC 2019, Div. 9, Part 4, Ch. 1, Sec. 16002.

regional office come together and do their best to find placements that are suitable for the children.

Another characteristic feature of how children are assigned placements in the field, is that the process is done as quickly as possible. In most cases, children must be placed on very short notice. Furthermore, even if a social worker knows that a child will be removed in the near future (usually not more than a few days), a placement cannot be assigned until the child has been removed. The reason for this is precisely because foster homes are scarce and there are children in need of placements constantly. Therefore, social workers cannot hold placements and wait for children to be removed from home. It would mean that other children are not being placed, which social workers try to avoid as best they can.[9]

2.3 Data Description and Summary Statistics

The data used in this study comes from confidential county records of DCFS. The database used for the analysis includes the record of every child that was placed in a foster home at any point between January 1, 2011, and February 28, 2011, in LA County.[10] During this period, 2,087 children where assigned to a foster home at least once in LA County, and 2,358 placements were assigned in total across the nineteen regional offices in LA County. On average, roughly 40 placements are assigned everyday throughout the county. Table 1 contains summary statistics of the placements in the dataset.

2.4 Modeling Strategy

In what follows, I develop an econometric model with the objective of analyzing the determinants underlying placement assignment. The main focus is on placements that were assigned on the same day in the same regional office. That is, the model aims to explain what drives the matching between children and foster homes

[9] Children who enter foster care at times when there are no placements available are usually placed in Emergency-Foster Care or Emergency Shelter Care while a non-emergency placement can be found (usually a few days at most). Emergency placements are available 24/7, but are not suitable for stays lasting more than a few days.

[10] The confidentiality waiver needed to access the data granted access to a larger time period. However, I restrict the sample period to a two-month period for computational considerations. As it shall be seen in the coming sections, the econometric framework I develop in this paper is computationally intensive.

Table 1: Summary Statistics

	(1) n	(2) mean	(3) sd	(4) p5	(5) p25	(6) p50	(7) p75	(8) p95
Termination Reasons								
Disruption	2358	0.5093	0.5	0	0	1	1	1
Exit	2358	0.4237	0.4942	0	0	0	1	1
Emancipation	2358	0.05174	0.2215	0	0	0	0	1
Censored	2358	0.01527	0.1226	0	0	0	0	0
Duration								
Duration (days)	2358	255.4	343.9	5	35	131.5	339	898.4
Duration | Disrup	1201	164.6	242.7	4	22	74	186	623
Duration | Exit	999	304.1	304.8	5.45	66	223	439.2	879
Duration | Emanc	122	394.7	437.4	8.6	95	232	502	1400
Duration | Cens	36	1461	850.5	25.1	344.5	1969	1988	2002
Children Characteristics								
Time Since Removal (days)	2358	387.7	937.6	0	0	32	292	2184
Placement # In Spell	2358	2.75	2.582	1	1	2	3	8
Spell # in Child	2358	1.194	0.4626	1	1	1	1	2
Zero Waiting Time	2358	0.8562	0.3509	0	1	1	1	1
Waiting Time (days)	2358	0.9326	3.148	0	0	0	0	10.6
Age	2358	8.694	5.967	0.2037	2.916	8.485	14.54	17.35
Male	2358	0.4576	0.4983	0	0	0	1	1
Black	2358	0.3138	0.4641	0	0	0	1	1
Hispanic	2358	0.5424	0.4983	0	0	1	1	1
White	2358	0.1175	0.3221	0	0	0	0	1
Other Race	2358	0.02629	0.16	0	0	0	0	0
English	2358	0.8223	0.3823	0	1	1	1	1
Spanish	2358	0.1773	0.382	0	0	0	0	1
Other Language	2358	0.0004241	0.02059	0	0	0	0	0
Absence/Incapacitation	2358	0.2693	0.4437	0	0	0	1	1
Abuse/Severe Neglect	2358	0.2498	0.433	0	0	0	0	1
General Neglect	2358	0.4597	0.4985	0	0	0	1	1
Other Removal Reason	2358	0.0212	0.1441	0	0	0	0	0
Children Characteristics								
County Foster Home	2358	0.08567	0.2799	0	0	0	0	1
Agency Foster Home	2358	0.4258	0.4946	0	0	0	1	1
Group Home	2358	0.1158	0.32	0	0	0	0	1
Relative Home	2358	0.3728	0.4836	0	0	0	1	1
Distance Plac-Office (mi.)	2358	22.93	21.27	2.22	7.716	16.05	30.69	71.15
Distance Plac-School (mi.)	2358	18.13	23.77	0	0	7.983	26.9	72.73
No School	2358	0.2472	0.4315	0	0	0	0	1

Note: Summary statistics of placement outcomes and characteristics. The distance measures are at the zip-code level (foster home and school). They were computed using the Google Maps API. No School refers to children for which the dataset includes no school zip-code (presumably because the child does not go to school, or the data is missing). sd = standard deviation; p# refers to the #th percentile.

in cases in which two or more placements where assigned in the same day in the same regional office. For this purpose, I slice the data of placements into markets accordingly. The division of the data into markets also incorporates placements with relatives. That is, if a child was placed with a relative, I form an independent market consisting of a single child and a single home in which the assignment problem is trivial. The reason I keep "singleton" markets (i.e., with a single child and single home) is to study their outcomes.

The way I model placement assignment is by considering a single matchmaker that assigns placements in terms of their expected outcomes. That is, when there are several ways in which children and foster homes can be matched, the matchmaker is assumed to consider the expected outcomes of all prospective placements, and weigh them according to a specific utility function. I rationalize the observed matching by considering it as the optimal matching from the matchmaker's perspective. Apart from considering the expected outcomes of prospective placements, the matchmaker's problem also allows for children and foster homes to have idiosyncratic tastes for the type of foster home and child with whom they are matched. The model is designed to include the most prominent institutional features of foster care placement. That being said, the only feature I abstract away from in this paper is the placement of siblings. In what follows, I ignore the existence of siblings in the system, and focus on one-to-one matchings. The analysis of placement assignment with siblings is ripe ground for future research.

3 Model

3.1 Market of Foster Care Services

A market is a tuple $(C, H, \mathbf{X}, \mathbf{Y})$, where C is the set of available children; H is the set of available foster homes; $\mathbf{X} = (\mathbf{x}_c)_{c \in C}$ is the matrix of children's (observable) characteristics, i.e., $\mathbf{x}_c \in \mathcal{X} \subseteq \mathbb{R}^{\dim(\mathbf{x})}$ is the vector of characteristics of child $c \in C$, and $\mathbf{Y} = (\mathbf{y}_h)_{h \in H}$ is the matrix of the (observable) characteristics of available homes, i.e., $\mathbf{y}_h \in \mathcal{Y} \subseteq \mathbb{R}^{\dim(\mathbf{y})}$ is the vector of characteristics of home $h \in H$. In order to incorporate idiosyncratic preferences over children's and foster home's characteristics, I define *types* as a coarsening of characteristics. Let $X = \{x\}$ and $Y = \{y\}$ be the sets of child- and home-types; formally, they are finite partitions of $\mathcal{X}$ and $\mathcal{Y}$. Similarly, let $x_c \in X$ and $y_h \in Y$ denote the types of $c \in C$ and $h \in H$, respectively.

A one-to-one matching between children and foster homes is an indicator function $M : C \times H \to \{0, 1\}$ such that $\sum_{h \in H} M(c, h) \leq 1$ for all $c \in C$, and $\sum_{c \in C} M(c, h) \leq 1$ for all $h \in H$. That is, $M(c, h) = 1$ if child c is matched with home h, and 0 otherwise. For simplicity, I also write $(c, h) \in M$ if $M(c, h) = 1$. Let $\mathbb{M}(C, H)$ denote the set of feasible one-to-one matchings between C and H.

Matching a child and a home forms a placement. The outcome of a placement is given by $(T, R) \in \mathbb{R}_+ \times \mathcal{R}$, where T denotes the placement's duration, and R its termination reason. A placement may terminate because it is disrupted (d), the child exits to permanency (ex), or is emancipated (em). The set of termination reasons is thus $\mathcal{R} \equiv \{d, ex, em\}$. It is convenient to differentiate emancipation from the other termination reasons because the time to emancipation, denoted by T_{em}, is not random. I define the set of termination reasons with non-degenerate duration as $\mathcal{R}_0 = \{d, ex\}$.

Children are matched to foster homes on a daily basis within regional offices throughout the county. The unit of observation is a market, indexed by $i = 1, \ldots, n$. Markets correspond to office-days, and also incorporate the restriction that children need to be matched with their relatives whenever possible. That is, children for whom relatives are available as prospective foster parents have their own markets (consisting of a single child and a single foster home). The data consists on (1) a sample of markets, $(C_i, H_i, \mathbf{X}_i, \mathbf{Y}_i)_{i=1}^n$; (2) the matching chosen in each market, $\mathbf{M} = (M_i)_{i=1}^n$, where $M_i \in \mathbb{M}(C_i, H_i)$ for $i = 1, \ldots, n$, and (3) the outcomes of the assigned placements, $(\mathbf{T}_i, \mathbf{R}_i)_{i=1}^n$, where $\mathbf{T}_i = (T_{ch})_{(c,h) \in M_i}$, and $\mathbf{R}_i = (R_{ch})_{(c,h) \in M_i}$.

I take the data of markets, $(C_i, H_i, \mathbf{X}_i, \mathbf{Y}_i)_{i=1}^n$, as given (i.e., as exogenous variables). The observed matching and the realized outcomes, $(M_i, \mathbf{T}_i, \mathbf{R}_i)_{i=1}^n$, are the outcome (or endogenous) variables of the model. Note that this implies that there are no spillovers across office-days. Every day, in every office, a matching is assigned between the available children and foster homes taking the market as given. I outline the data generating process of the endogenous variables, $(M, \mathbf{T}, \mathbf{R})$, in the following sections.

3.2 Placement Assignment

Placements are assigned by a single (or representative) utilitarian matchmaker, who has preferences over realized outcomes $(T, R) \in \mathbb{R}_+ \times \mathcal{R}$. The matchmaker's preferences are represented by the utility function:

$$u(T, R; T_{em}) = \mu_R + \varphi_R \log T + \bar{\varphi}_R \log T_{em}, \tag{1}$$

where $\mu_R, \varphi_R, \bar{\varphi}_R \in \mathbb{R}$, are unknown parameters for $R \in \mathcal{R}$. The parameter μ_R measures the preference over termination reason $R \in \mathcal{R}$, regardless of duration; φ_R is the marginal utility of duration conditional on terminating due to $R \in \mathcal{R}$. The utility function also includes the time to emancipation in its third term to control for the fact that placements involving younger children may have ex-ante longer durations. For example, if $\bar{\varphi}_R = -\varphi_R$, the matchmaker cares about duration relative to the time to emancipation. More generally, one can see that the sign of the marginal rate of substitution between duration and age, conditional on termination reason $R \in \mathcal{R}$, is equal to the sign of $\varphi_R/\bar{\varphi}_R$.

Consider a prospective placement $(c, h) \in C \times H$. Let $\boldsymbol{\mathcal{I}}_{ch}$ denote the information that the matchmaker has on its outcome distribution. The total payoff of placing child $c \in C$ in home $h \in H$ to the matchmaker is given by:

$$V(c, h) = \pi(c, h) + \varepsilon_{cy_h} + \eta_{x_c h}, \tag{2}$$

where $\pi(c, h) := \mathbb{E}\left[u(\tilde{T}, \tilde{R}; T_{em}) \,|\, \boldsymbol{\mathcal{I}}_{ch}\right]$ captures the preferences and information available to the matchmaker about the placement's outcome. I specify the distribution of $(\tilde{T}, \tilde{R}) \,|\, \boldsymbol{\mathcal{I}}_{ch}$ in the next section.[11] The remaining two terms in (2), ε_{cy_h} and $\eta_{x_c h}$, capture idiosyncratic taste variation across children and foster homes (which is unobservable to the econometrician). Specifically, ε_{cy} captures the payoff of matching child c with a home of type $y \in Y$, and η_{xh} that of matching home h with a child of type $x \in X$. In this sense, the model incorporates the preferences that children may have for being placed in specific types of homes, and those of homes for taking care of particular types of children. More generally, the taste variation terms are aimed to capture type-specific idiosyncratic unobservables that affect placement

[11] I differentiate random variables that are observable to the econometrician from their realized values with a tilde; $(\tilde{T}, \tilde{R})$ denotes the unrealized (random) placement outcome, while $(T, R) \in \mathbb{R}_+ \times \mathcal{R}$ denotes its realization.

assignment (e.g., the matchmaker may also have preferences over forming certain types of placements, regardless of their outcomes).

The matchmaker chooses the matching $M \in \mathbb{M}(C, H)$ that maximizes its aggregate payoff. Since $V(c, h)$ is observable to the matchmaker for all $(c, h) \in C \times H$, the observed matching is the solution to the following linear programming problem:

$$(3) \qquad \max \left\{ \sum_{c \in C, h \in H} M(c, h) V(c, h) : M \in \mathbb{M}(C, H) \right\}.$$

I restrict attention to matchings in which no child is left unmatched while there is an unmatched home. That is, besides incorporating the natural constraints that every child can be matched with at most one home (and vice versa), the set of feasible matchings $\mathbb{M}(C, H)$ satisfies:

$$(4) \qquad M \in \mathbb{M}(C, H) \quad \Leftrightarrow \quad \sum_{c \in C, h \in H} M(c, h) = \min\{|C|, |H|\}.$$

3.3 Placement Outcomes

Prospective placements are indexed by $(c, h) \in C \times H$. For simplicity, consider a generic placement and omit such index in this section. The full vector of characteristics of a placement is given by $\boldsymbol{\mathcal{I}} = (\mathbf{x}, \mathbf{y}, \boldsymbol{\omega})$, where $(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}$ are the observable child- and home-characteristics, and $\boldsymbol{\omega} \in \mathbb{R}^{\dim(\boldsymbol{\omega})}$ is a vector of characteristics *not* observed by the econometrician. The distribution of a placement's outcome, $(\tilde{T}, \tilde{R})$, depends on its full vector of characteristics, $\boldsymbol{\mathcal{I}}$.

I model placement outcomes as the result of mixed competing risks. Consider a generic placement with characteristics $\boldsymbol{\mathcal{I}} = (\mathbf{x}, \mathbf{y}, \boldsymbol{\omega})$. Let $\tilde{T}_R$ be the latent duration associated to the "risk" of terminating due to reason $R \in \mathcal{R}_0$. Up to censoring, due to the sample period or emancipation, a placement's outcome is determined by the least latent duration. Denote the time to the end of the sample period by T_{cen}, and indicate censored placements by $R = cen$. To simplify notation, I define $\tilde{T}_{em} = T_{em}$ and $\tilde{T}_{cen} = T_{cen}$ as the degenerate latent durations corresponding to the time to emancipation and the end of the sample period, respectively. Formally, the outcome of a placement is given by:

$$(5) \qquad \tilde{T} = \min\left\{\tilde{T}_R : R \in \mathcal{R} \cup \{cen\}\right\}, \quad \text{and} \quad \tilde{R} = \arg\min\left\{\tilde{T}_R : R \in \mathcal{R} \cup \{cen\}\right\}.$$

Under the above specification, a placement is emancipated (or censored) if and only if it has not been disrupted or has exited to permanency by its emancipation date (or the end of the sample period). Note that each placement in the data is subject to either emancipation or censoring due to the sample period, depending on which of T_{em} and T_{cen} is lower. Both types of censoring, due to emancipation and the end of the sample period, are equivalent in terms of the likelihood of the latent durations. However, they are not equivalent from the matchmaker perspective, who has a preference over the emancipation likelihood and the time to emancipation. Censoring due to the sample period is only statistical in nature.

Assumption 1 (Unobserved heterogeneity)**.** *The unobservable characteristics of a placement are given by the vector $\boldsymbol{\omega} = (\omega_R)_{R\in\mathcal{R}_0}$. Furthermore,*

$$(6) \qquad \boldsymbol{\omega} \sim N(0, \boldsymbol{\Sigma}_\omega),$$

where $\boldsymbol{\Sigma}_\omega$ is a positive semidefinite and symmetric matrix of size $|\mathcal{R}_0| \times |\mathcal{R}_0|$.

Assumption 2 (Burr hazards)**.** *Conditional on a placement's characteristics, $\boldsymbol{\mathcal{I}}$, the latent durations, $\{T_R : R \in \mathcal{R}_0\}$, are independent. Furthermore, the conditional distribution of T_R is determined by the following Burr hazard rate*[12]*,*

$$(7) \qquad \lambda_R(T\,|\,\boldsymbol{\mathcal{I}}) = \frac{k_R(\boldsymbol{\mathcal{I}})\alpha_R T^{\alpha_R-1}}{1+\gamma_R^2 k_R(\boldsymbol{\mathcal{I}})T^{\alpha_R}}, \quad R \in \mathcal{R}_0,$$

where $k_R(\boldsymbol{\mathcal{I}}) \equiv \exp\{\omega_R + g(\mathbf{x},\mathbf{y})\beta_R\}$ with $\beta_R \in \mathbb{R}^{\dim(\beta)}$, $g : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}^{\dim(\beta)}$, $\alpha_R > 0$, and $\gamma_R \geq 0$.

Assumption 2 specifies the distribution of placement outcomes from the perspective of the matchmaker, $(\tilde{T},\tilde{R})\,|\,\boldsymbol{\mathcal{I}}$. The matchmaker's additional information, $\boldsymbol{\omega}$, consists of unobservable frailty terms, $(\omega_R)_{R\in\mathcal{R}_0}$, which may shift the hazard rate associated to each "risk" (termination reason) upwards or downwards. Since such frailty terms are not observable to the econometrician, the distribution $(\tilde{T},\tilde{R})\,|\,\boldsymbol{\mathcal{I}}$ is

[12]The hazard rate of the random variable $\tilde{T}$ is the function defined by $\lambda(T) = f(T)/\bar{F}(T)$, where f denotes the probability density function of $\tilde{T}$, and $\bar{F}$ its survivor function. (The survivor function is defined by $\bar{F}(T) \equiv 1 - F(T)$, where F is the variable's cumulative distribution function.)

not observed directly in the data. The outcome distribution is "mixed" by the distribution of the unobservable frailty terms; one must integrate out $\boldsymbol{\omega}$ to recover the distribution of outcomes in the data. However, note that the distribution of $\boldsymbol{\omega}$ across the placements observed in the data is *not* equal to the unconditional distribution specified in Assumption 1. The distribution of $\boldsymbol{\omega}$ across the placements in the data is conditional on being matched, i.e., to that of $\boldsymbol{\omega}_{ch} \,|\, M(c,h) = 1$.

The Burr specification in Assumption 2 is a standard parametric assumption used in duration models (e.g. Lancaster 1990; Wooldridge 2010).[13] The Burr distribution has the main advantage of being flexible yet tractable. It generalizes other well-known duration distributions, such as the Exponential ($\gamma_R = 0, \alpha_R = 1$), Weibull ($\gamma_R = 0$), and Log-Logistic ($\gamma_R = 1$). A convenient feature of this distribution is that its integrated hazard rate has a closed form, and hence, also its survivor function and likelihood. The parameters α_R and γ_R govern the duration-dependence of the hazard function, which may be flat, monotonic (positive or negative), or have an inverse-U shape. The function g is a shorthand for the covariates used in the model, all of which are derived from observable characteristics. Besides including stand-alone covariates, $g(\mathbf{x}, \mathbf{y})$ may include interactions between variables in $\mathbf{x}$ and $\mathbf{y}$, and other non-trivial transformations, such as distance measures. The effect of the covariates on each hazard rate is controlled by the coefficients in β_R. Since the function λ_R is monotonic in k_R, the sign of the coefficients in β_R indicate the direction in which the covariates shift the hazard rates. A higher hazard rate, say λ_R, implies that a placement is more likely to terminate sooner and due to termination reason $R \in \mathcal{R}_0$.

Assumption 1 specifies the joint distribution of $\boldsymbol{\omega} = (\omega_R)_{R \in \mathcal{R}_0}$ up to the unknown covariance matrix $\boldsymbol{\Sigma}_\omega$. Assuming that $\boldsymbol{\omega}$ has zero mean is without loss of generality, as long as the covariates in the hazard function include a constant. Intuitively, the covariance matrix $\boldsymbol{\Sigma}_\omega$ captures the extent of the variation in the observed outcomes not captured through placement characteristics. Moreover, the correlation between the individual frailty terms introduces dependence among the latent durations. Such correlation captures, for example, if (a) children who are less likely to reach permanency are also more likely to experience disruptions (because, say, they experienced worse conditions during their upbringing, and this has an impact in their current behavior), or (b) children who are more likely to exit the system

[13] Another common application of the Burr distribution, also known as the Singh-Maddala distribution, is to model the distribution of income (Singh and Maddala 1976).

sooner are also more likely to experience disruptions (because, say, foster parents are less invested in nurturing long and stable relationships with children who will leave their households sooner).

Collect the parameters of the hazard rates in $\boldsymbol{\alpha} = (\alpha_R)_{R\in\mathcal{R}_0}$, $\boldsymbol{\gamma} = (\gamma_R)_{R\in\mathcal{R}_0}$, and $\boldsymbol{\beta} = (\beta_R)_{R\in\mathcal{R}_0}$. The conditional outcome distribution, $(\tilde{T}, \tilde{R})\,|\,\boldsymbol{\mathcal{I}}$, is fully specified in Assumption 2 up to the unknown vector of parameters

$$\boldsymbol{\theta}_T \equiv (\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}).$$

3.4 Observed Matching

In this section, I consider a generic market $(C, H, \mathbf{X}, \mathbf{Y})$, and omit its index $i = 1, \ldots, n$ for simplicity. The problem of the matchmaker in (3) is a deterministic problem over matchings. However, from the econometrician's perspective, the observed matching is the realization of a random variable since $V(c, h)$ is not fully observable. Specifically, an econometrician does not observe the frailty terms $(\boldsymbol{\omega}_{ch})_{(c,h)\in C\times H}$, or the taste variation terms, $\varepsilon_c = (\varepsilon_{cy})_{y\in Y}$ for every $c \in C$, and $\eta_h = (\eta_{xh})_{x\in X}$ for every $h \in H$.

Assumption 3 (Multinomial Probit)**.** *The taste variation terms are independent multivariate normal random vectors. Namely,*

$$\varepsilon_c \sim N(0, \boldsymbol{\Sigma}_\varepsilon), \quad \textit{and} \quad \eta_h \sim N(0, \boldsymbol{\Sigma}_\eta), \tag{8}$$

where $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$ are positive semidefinite and symmetric matrices. Their sizes are $|Y|\times|Y|$ and $|X| \times |X|$, respectively. Furthermore, $\varepsilon_c \perp \varepsilon_{c'}$ for all $c, c' \in C$, and $\eta_h \perp \eta_{h'}$ for all $h, h' \in H$. Also, ε_c, η_h, and $\boldsymbol{\omega}_{c'h'}$ are mutually independent for all $(c, h), (c', h') \in C \times H$.

Under Assumption 3, the observed matching is a realization of the following random variable:

$$\tilde{M}(C, H, \mathbf{X}, \mathbf{Y}) = \arg\max \left\{ \sum_{c\in C, h\in H} M(c, h)\pi(c, h) + \upsilon_M : M \in \mathbb{M}(C, H) \right\}, \tag{9}$$

where υ_M is the composite error term given by

$$\upsilon_M \equiv \sum_{c\in C, h\in H} M(c,h)[\varepsilon_{cy_h} + \eta_{x_c h}]. \tag{10}$$

Since the composite error term, υ_M, follows a multivariate normal distribution, the matching problem takes the form of a mixed multinomial probit. Below , I show that the distribution of the individual taste variation parameters, ε_c and η_y, can be backed out from the distribution of the composite error term υ_M. Therefore, the distribution of the taste variation parameters can be obtained directly from the matching data.

Assumption 3 also includes several independence assumptions. First, the unobservable taste variation components are independent across parties, i.e., $\varepsilon_c \perp \varepsilon_{c'}$, $\eta_h \perp \eta_{h'}$, and $\varepsilon_c \perp \eta_h$. This assumption rules out unobservable interdependencies among placement assignments by considering preferences over types as independent across children and foster homes. Second, the unobservable frailty terms are independent across placements, i.e., $\boldsymbol{\omega}_{ch} \perp \boldsymbol{\omega}_{c'h'}$. This assumption rules out unobservable interdependencies among placement outcomes. Conditional on being matched, the outcome of (c,h) is independent of that of (c',h'). Third, the taste variation terms, ε_c and η_h, are independent of the frailty terms in $\boldsymbol{\omega}_{ch}$. This assumption separates the unobservables affecting placement assignments into two groups. On the one hand, $\boldsymbol{\omega}_{ch}$ contains unobservables that affect placement assignments through their expected outcomes (i.e., outcome-relevant unobservables). On the other hand, ε_c and η_h capture the rest of the unobservables which affect the matching but are independent of outcomes.

Collect the preference parameters in $\boldsymbol{\mu} = (\mu_R)_{R\in\mathcal{R}_0}$, $\boldsymbol{\varphi} = (\varphi_R)_{R\in\mathcal{R}_0}$, $\bar{\boldsymbol{\varphi}} = (\bar{\varphi}_R)_{R\in\mathcal{R}_0}$, $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_\varepsilon, \boldsymbol{\Sigma}_\eta)$, and define

$$\boldsymbol{\theta}_M \equiv (\boldsymbol{\mu}, \boldsymbol{\varphi}, \bar{\boldsymbol{\varphi}}, \boldsymbol{\Sigma}).$$

3.5 Two-by-two Example

In this section, I consider a simple example to illustrate how the model allows for the matching observed in the data to depend on distinct factors. Consider a market with two children and two homes, let $C = \{c_1, c_2\}$ and $H = \{h_1, h_2\}$. Let x_1 and x_2

denote the types of children c_1 and c_2, respectively, and y_1 and y_2, those of homes h_1 and h_2. Let $\varepsilon_1 = (\varepsilon_{11}, \varepsilon_{12})$ and $\varepsilon_2 = (\varepsilon_{21}, \varepsilon_{22})$ be the unobservable tastes of child c_1 and c_2 for home-types y_1 and y_2, respectively, where I take the liberty of writing $\varepsilon_{kj} \equiv \varepsilon_{c_k y_j}$. Similarly, let $\eta_1 = (\eta_{11}, \eta_{21})$ and $\eta_2 = (\eta_{12}, \eta_{22})$ be the unobservable tastes of homes h_1 and h_2 for child-types x_1 and x_2, respectively, with $\eta_{kj} \equiv \eta_{x_k h_j}$. Finally, let $(\boldsymbol{\omega}_{11}, \boldsymbol{\omega}_{12}, \boldsymbol{\omega}_{21}, \boldsymbol{\omega}_{22})$ be the unobservable vectors of frailty terms of each prospective placement, i.e., $\boldsymbol{\omega}_{kj} \equiv (\omega_{c_k h_j, R})_{R \in \mathcal{R}_0}$ for $j, k = 1, 2$.

Let $\pi_{kj} \equiv \pi(c_k, h_j)$ denote the payoff of assigning each prospective placement, which is a function of $\boldsymbol{\omega}_{kj}$, for $j, k = 1, 2$. The set of feasible matchings $\mathbb{M}(C, H)$ contains two matchings: M, which assigns placements (c_1, h_1) and (c_2, h_2), and M', which assigns placements (c_1, h_2) and (c_2, h_1). Let $\mathcal{V}$ and $\mathcal{V}'$ denote their respective aggregate payoffs, i.e.,

$$\mathcal{V} = V(c_1, h_1) + V(c_2, h_2) = (\pi_{11} + \varepsilon_{11} + \eta_{11}) + (\pi_{22} + \varepsilon_{22} + \eta_{22}) \tag{11}$$

$$\mathcal{V}' = V(c_1, h_2) + V(c_2, h_1) = (\pi_{12} + \varepsilon_{12} + \eta_{12}) + (\pi_{21} + \varepsilon_{21} + \eta_{21}) \tag{12}$$

Matching M is chosen over M' if and only if $\mathcal{V} \geq \mathcal{V}'$ (the event $\mathcal{V} = \mathcal{V}'$ has zero probability). In principle, all the terms in (11) and (12) might differ, implying that observing matching M over M' might result for numerous reasons, e.g., the expected outcome of placement (c_1, h_2) or (c_2, h_1) is unfavorable relative to that of (c_1, h_1) or (c_2, h_2) (i.e., π_{12} or π_{21} are low relative to π_{11} or π_{22}) . Alternatively, child c_k might have a higher than usual preference for being matched with a home of type y_k (i.e., ε_{11} or ε_{22} are particularly high), or home h_j might have a higher than usual preference for being matched with a child of type x_j (i.e., η_{11} or η_{22} are high relative to η_{12} or η_{21}).

Now consider the case in which $y_1 = y_2$, so that $\varepsilon_{11} = \varepsilon_{12}$ and $\varepsilon_{21} = \varepsilon_{22}$. In such case, matching M is chosen over M' if and only if

$$(\pi_{11} + \eta_{11}) + (\pi_{22} + \eta_{22}) \geq (\pi_{12} + \eta_{12}) + (\pi_{21} + \eta_{21})\,. \tag{13}$$

In this case, even though the unobservable taste terms of both children may differ, i.e., $\varepsilon_{11} \neq \varepsilon_{21}$, the preferences of children over home-types play no role in the determination of the optimal matching. Similarly, if the children are also of the same type, $x_1 = x_2$, then matching M is chosen over M' if and only if

$$\pi_{11} + \pi_{22} \geq \pi_{12} + \pi_{21}. \tag{14}$$

In this case, the optimal matching is determined only on the basis of expected outcomes. Importantly, the event $\mathcal{V} \geq \mathcal{V}'$ is still random from the econometrician's perspective, since (14) depends on the unobservable frailty terms, $\boldsymbol{\omega}_{kj}$ for $j, k = 1, 2$.

3.6 Expected Placement Outcomes

In this section I show in more detail how the payoff function depends on a placement's expected outcomes. Using (1) and the definition of $\pi(c, h)$, we obtain

$$\pi(c,h) = \sum_{R\in\mathcal{R}} \mathbb{P}(\tilde{R} = R|\boldsymbol{\mathcal{I}}_{ch}) \left\{\mu_R + \varphi_R \mathbb{E}\left[\log \tilde{T}\,|\,\tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch}\right] + \bar{\varphi}_R \log T_{em,c}\right\}. \tag{15}$$

Therefore, the expected placement outcomes that are relevant for the matchmaker's payoff are the termination probability, $\mathbb{P}(\tilde{R} = R|\boldsymbol{\mathcal{I}}_{ch})$, and the conditional expected log-duration, $\mathbb{E}\left[\log \tilde{T}\,|\,\tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch}\right]$, of each termination reason $R \in \mathcal{R}$. The expected placement outcomes can be computed using standard results in survival analysis (e.g. Lancaster 1990; Kalbfleisch and Prentice 2002).[14] Namely, for $R \in \mathcal{R}_0$,

$$\mathbb{P}(\tilde{R} = R|\boldsymbol{\mathcal{I}}_{ch}) = \int_0^{T_{em,c}} \bar{F}(T\,|\boldsymbol{\mathcal{I}}_{ch})\lambda_R(T\,|\boldsymbol{\mathcal{I}}_{ch})dT \tag{16}$$

$$\mathbb{E}\left[\log \tilde{T}\,|\,\tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch}\right] = \int_0^{T_{em,c}} \log T \left[\frac{\bar{F}(T\,|\boldsymbol{\mathcal{I}}_{ch})\lambda_R(T\,|\boldsymbol{\mathcal{I}}_{ch}}{\mathbb{P}(\tilde{R} = R|\boldsymbol{\mathcal{I}}_{ch})}\right] dT, \tag{17}$$

where $\bar{F}(T\,|\boldsymbol{\mathcal{I}}_{ch})$ denotes the conditional survival function of $\tilde{T}$, given by

$$\bar{F}(T\,|\boldsymbol{\mathcal{I}}_{ch}) = \exp\left\{-\sum_{R\in\mathcal{R}_0} \gamma_R^{-2} \log\left[1 + \gamma_R^2 k_R(\boldsymbol{\mathcal{I}}_{ch})T^{\alpha_R}\right]\right\}. \tag{18}$$

[14] To observe why (16) holds, it suffices to note that $\bar{F}(T\,|\,\boldsymbol{\mathcal{I}}_{ch})\lambda_R(T\,|\,\boldsymbol{\mathcal{I}}_{ch})$ is the likelihood of the placement having duration T and terminating due to $R \in \mathcal{R}_0$. The probability of terminating due to $R \in \mathcal{R}_0$ is the integral of this likelihood over the support of $\tilde{T}$, $[0,\ T_{em,c}]$. Similarly, to observe why (17) holds, it suffices to note that the quotient in brackets in (17) is the probability density function (pdf) of $\tilde{T}\,|\,\tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch}$. To see this, note that the likelihood of the event $(\tilde{T}, \tilde{R}) = (T, R)$ may also be written as $\mathbb{P}(\tilde{R} = R\,|\,\boldsymbol{\mathcal{I}}_{ch})f(T\,|\,\tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch})$, where $f(T\,|\,\tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch})$ denotes the pdf of $\tilde{T}\,|\,\tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch}$. Expression (18) also follows from standard results. Namely, the survivor function of the duration in a competing risks model is given by $\bar{F}(T) = \exp\left\{-\sum_{R\in\mathcal{R}_0}\int_0^T \lambda_R(S)dS\right\}$.

Simple calculations show that the resulting integrals in (16) and (17) have no closed-form.[15] Therefore, to compute the payoff function of placement $(c, h) \in C \times H$, one needs to compute the integrals in (16) and (17) numerically at $\boldsymbol{\mathcal{I}}_{ch} = (\mathbf{x}_c, \mathbf{y}_h, \boldsymbol{\omega}_{ch})$, obtain the expected placements outcomes, and replace the respective values in (15).

In order to observe how the aggregate payoff of matching $M \in \mathbb{M}(C, H)$ depends on the expected placement outcomes of the assigned placements, note that

$$
\begin{aligned}
\text{(19)} \qquad \sum_{c,h} M(c,h)\pi(c,h) = \sum_{R\in\mathcal{R}} \Bigg\{ & \left[\sum_{c,h} M(c,h)\mathbb{P}(\tilde{R} = R \,|\, \boldsymbol{\mathcal{I}}_{ch}) \right] \mu_R \\
& + \left[\sum_{c,h} M(c,h)\mathbb{P}(\tilde{R} = R \,|\, \boldsymbol{\mathcal{I}}_{ch})\mathbb{E}\left(\log \tilde{T} \,|\, \tilde{R} = R, \boldsymbol{\mathcal{I}}_{ch} \right) \right] \varphi_R \\
& + \left[\sum_{c,h} M(c,h)\mathbb{P}(\tilde{R} = R \,|\, \boldsymbol{\mathcal{I}}_{ch}) \log T_{em,c} \right] \bar{\varphi}_R \Bigg\},
\end{aligned}
$$

where the sums are over $c \in C, h \in H$. Hence, conditional on the matchmaker's information on every prospective placement, $(\boldsymbol{\mathcal{I}}_{ch})_{(c,h)\in C\times H}$, the problem of the matchmaker in (9) takes the form of a multinomial probit. The "systematic" or "observed" portion of the aggregate payoff of matching $M \in \mathbb{M}(C, H)$, given in (19), is a linear index on the parameters of the matchmaker's utility function, $(\boldsymbol{\mu}, \boldsymbol{\varphi}, \bar{\boldsymbol{\varphi}})$. The "covariates" of such linear index are sums of the expected outcomes of all the assigned placements under M, which, in essence, are non-linear transformations of the covariates of the assigned placements, $\{g(\mathbf{x}_c, \mathbf{y}_h) : M(c, h) = 1\}$. The unconditional problem of the matchmaker takes the form of a mixed multinomial probit since one must integrate out the unobservable part of $(\boldsymbol{\mathcal{I}}_{ch})_{(c,h)\in C\times H}$, i.e., $(\boldsymbol{\omega}_{ch})_{(c,h)\in C\times H}$.

[15] The fact that these integrals have no closed-form is a common feature among most commonly used duration distributions. A notable exception, perhaps the only one, is the competing risks model with symmetric Weibull hazards (all hazards have the same shape parameter). In our case, this corresponds to the case with $\gamma_R = 0$ and $\alpha_R = \alpha$ for all $R \in \mathcal{R}_0$. In such case, the termination probabilities have the same form as the choice probabilities of the multinomial logit, and are constant across time. As shall be seen in next sections, this specification, although attractive for its computational tractability, is too restrictive for the present case.

4 Identification

4.1 Outcome Distribution

Absent matching, the data on observed outcomes is sufficient to identify the parameters of the distribution of outcomes, $(\mathbf{\Sigma}_\omega, \boldsymbol{\theta}_T)$. This observation follows from Heckman and Honoré (1989), who show that the joint distribution of the latent durations in a competing risks model is non-parametrically identified as long as (1) the model includes covariates; (2) the hazard rates of the latent durations have at least one common covariate with a different coefficient in each hazard rate; (3) such covariate is continuous and unbounded, and (4) the mixing distribution is sufficiently smooth (and satisfies certain regularity conditions at the limit). All of these conditions are met given Assumptions (1) and (2). The continuous and unbounded covariates are distance measures, e.g., the distance between children's schools and foster homes, which has a termination-specific coefficient.

Once we take into account the matching part of the model, one must recognize that the distribution of $\boldsymbol{\omega}$ across the placements observed in the data, in general, differs from the unconditional distribution specified in Assumption 1. The distribution of $\boldsymbol{\omega}$ across the placements observed in the data is given by $\boldsymbol{\omega}_{ch} \,|\, \tilde{M}(c,h) = 1$, where $\tilde{M}$ is the random variable defined in (9). Hence, the distribution of $\boldsymbol{\omega}$ observed in the data depends on all the variables involved in the matchmaker's problem. In order to identify this distribution, the model relies on the random variation on the exogenous variables $(C, H, \mathbf{X}, \mathbf{Y})$. The simplest way to see why this variation is sufficient to identify the parameters in the unconditional distribution of $\boldsymbol{\omega}$ is to consider placements in singleton markets. Note that the distribution of $\boldsymbol{\omega}$ for placements assigned in markets with $|C| = |H| = 1$ is the same as its unconditional distribution. That is, if $|C| = |H| = 1$, the matchmaker's problem is trivial, which implies that the event $\{M(c,h) = 1\}$ is uninformative, and the likelihood of such placement's outcome is the same as its unconditional one. More generally, in non-singleton markets, exogenous variation in $(C, H, \mathbf{X}, \mathbf{Y})$ identifies the unconditional mixing distribution in a similar way in which instruments are used in standard sample selection models (e.g. Heckman 1979). One needs exogenous variation that affects the likelihood of being "selected" (i.e., having an observable outcome) that is independent of the outcome itself.

Another aspect that differs from the standard competing risks framework is that the matching may induce endogeneity, which leads to bias when estimating the coefficients of the covariates in the hazard functions. This observation was first noted in the literature by Ackerberg and Botticini (2002) in a setting of contract choice. Their setup is different to the one here, but the underlying intuition is the same. In a reduced-form setting, they show that when the outcome of a match (in their case, a joint sharecropping contract) depends on the characteristics of both parties involved in the match, the presence of unobservables correlated with the matching and the outcome lead to endogeneity. The matching affects the joint distribution of a match's characteristics, causing them to become correlated with the error term in a regression. To see this in our case, write the latent duration as follows[16]

$$\log \tilde{T}_R = K_R - g(\mathbf{x}, \mathbf{y})\beta_R/\alpha_R - \omega_R/\alpha_R + error_R, \tag{20}$$

where $error_R \equiv \log \tilde{T}_R - \mathbb{E}\left[\log \tilde{T}_R \,|\, \boldsymbol{\mathcal{I}}\right]$ is an exogenous error term, and

$$K_R \equiv \alpha_R^{-1}\left[\psi(1) - \psi(\gamma_R^{-2}) + \log \gamma_R^{-2}\right] \tag{21}$$

is a constant (ψ denotes the digamma function, $\psi(x) \equiv d \log \Gamma(x)/dx$, where Γ is the gamma function). In (20), one can see how the covariates affect the latent log-durations in an analogous way to a linear regression. At first glance, the covariates in $g(\mathbf{x}, \mathbf{y})$ seem to be exogenous. The unconditional distributions of ω_R and $error_R$ are independent of $(\mathbf{x}, \mathbf{y})$. However, the joint distribution of $(\mathbf{x}, \mathbf{y})$ across the assigned placements, $(\mathbf{x}_c, \mathbf{y}_h) \,|\, \tilde{M}(c, h) = 1$, is determined by the matching. Note that $\mathbf{y}_h = \sum_{h' \in H} \tilde{M}(c, h')\mathbf{y}_{h'}$. Hence, the covariates derived from $\mathbf{y}$ are no longer independent from the error term $-\omega_R/\alpha_R + error_R$ in (20). A symmetric argument shows that the same holds for the covariates derived from $\mathbf{x}$.

To fix this endogeneity problem, Ackerberg and Botticini (2002) suggest using instrumental variables that affect the matching, but are independent of outcomes. In the present case, this exogenous variation comes through $(C, H, \mathbf{X}, \mathbf{Y})$. Two placements that are observationally equivalent, say (c, h) and (c', h') with $(\mathbf{x}_c, \mathbf{y}_h) = (\mathbf{x}_{c'}, \mathbf{y}_{h'})$, will not have the same mixing distribution if they are assigned in distinct markets. If (say) the first placement is assigned in market $(C, H, \mathbf{X}, \mathbf{Y})$, and the sec-

[16] Expression (20) is a well-known feature of the Burr distribution (Lancaster 1990). Indeed, the fact that the log-duration can be written in the form of a linear regression is a characteristic feature of all accelerated failure time models, the Burr duration model included.

ond one in $(C', H', \mathbf{X}', \mathbf{Y}')$, then the distribution of $\boldsymbol{\omega} \mid \tilde{M}(c,h) = 1$, in general, will be distinct to that of $\boldsymbol{\omega} \mid \tilde{M}'(c',h') = 1$. The matchings chosen in both markets, $\tilde{M}$ and $\tilde{M}'$, are independent random variables with distinct distributions. This identification strategy has been used in the contracting literature since the seminal contribution of Ackerberg and Botticini (2002) (e.g. Sørensen 2007; Ewens, Gorbenko, and Korteweg 2019).

4.2 Matching Distribution

The identification of the parameters in the matchmaker's utility function, $(\boldsymbol{\mu}, \boldsymbol{\varphi}, \bar{\boldsymbol{\varphi}})$, is straightforward once the mixing distribution is identified, and one sets $\varphi_{em} = 0$. Setting $\varphi_{em} = 0$ is necessary since the time to emancipation appears twice in $u(T, R; T_{em})$ for $R = em$, see (1). As mentioned above, see (19), the matching problem is a multinomial probit with index linear on $(\boldsymbol{\mu}, \boldsymbol{\varphi}, \bar{\boldsymbol{\varphi}})$.

Finally, I discuss the identification of the covariance matrices of the taste variation terms, $\boldsymbol{\Sigma}_{\varepsilon}$ and $\boldsymbol{\Sigma}_{\eta}$. Let $\sigma_{\varepsilon}(y, y')$ be the (y, y')-th entry of $\boldsymbol{\Sigma}_{\varepsilon}$, i.e., $\sigma_{\varepsilon}(y, y') = \text{cov}(\varepsilon_{cy}, \varepsilon_{cy'})$. Similarly, let $\sigma_{\eta}(x, x') = \text{cov}(\eta_{xh}, \eta_{x'h})$. From (10), note that the vector of composite error terms, $\boldsymbol{\upsilon} \equiv (\upsilon_M)_{M \in \mathbb{M}(C,H)}$, follows a zero-mean multivariate normal distribution with covariance structure given by (a detailed proof is given in Appendix B):

$$\text{cov}(\upsilon_M, \upsilon_{M'}) = \sum_{c \in C} \sigma_{\varepsilon}(y_{M(c)}, y_{M'(c)}) + \sum_{h \in H} \sigma_{\eta}(x_{M(h)}, x_{M'(h)}), \tag{22}$$

where I write $M(c) = h \Leftrightarrow M(h) = c \Leftrightarrow M(c,h) = 1$. To deal with unmatched children in (22), set $\sigma_{\varepsilon}(y_{M(c)}, y_{M'(c)}) \equiv 0$ if c is unmatched in either M or M'. Standard results in discrete choice models (e.g. Train 2009) show that the covariance matrix of $\boldsymbol{\upsilon}$ is identified up to location and scale normalizations.

Assumption 4 (Covariance Normalization). *There exists $x_0 \in X$ and $y_0 \in Y$ such that $\sigma_{\eta}(x_0, x) = 0$ for every $x \in X$, and $\sigma_{\varepsilon}(y_0, y_0) = 1$.*

Assumption 4 imposes the necessary normalizations to identify the covariance matrices $\boldsymbol{\Sigma}_{\varepsilon}$ and $\boldsymbol{\Sigma}_{\eta}$. First, it imposes a location normalization by assuming that there exists a child-type, x_0, for which the taste variation term of every home equals to zero. Second, a scale normalization is assumed by assuming there exists a home-type, y_0, for which the variance of the corresponding taste variation term equals one for every child.

Proposition 1. *Under Assumption 4, the covariance matrices $\mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\eta$ are identified.*

The proof of Proposition 1 is provided in Appendix B. The proof exploits that the distribution of the taste variation terms is the same regardless of the types of the other available children and homes in the market. The proof relies on analyzing the identified elements of the covariance matrix of $\boldsymbol{\upsilon}$ in specific markets with particular types of children and homes, and use the normalization in Assumption 4 and the covariance structure in (22) to back out the covariance matrices $\mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\eta$.

Collect all the parameters of the model in $\boldsymbol{\theta} = [\mathbf{\Sigma}_\omega, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M]$. Let $\mathbf{\Theta} \in \mathbb{R}^{\dim(\boldsymbol{\theta})}$ be the parameter space. That is, $\mathbf{\Theta}$ is the subset of $\mathbb{R}^{\dim(\boldsymbol{\theta})}$ that incorporates the following parameter restrictions: $\alpha_R > 0, \gamma_R \geq 0$ for every $R \in \mathcal{R}_0$, $\varphi_{em} = 0$, $\mathbf{\Sigma}_\eta$ such that $\sigma_\eta(x_0, x) = 0$ for every $x \in X$, $\mathbf{\Sigma}_\varepsilon$ such that $\sigma_\varepsilon(y_0, y_0) = 1$, and $\mathbf{\Sigma}_\omega$, $\mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\eta$ are positive semidefinite and symmetric matrices.

5 Estimation

In this section, I explain how to obtain a consistent, efficient, and asymptotically normal estimator of $\boldsymbol{\theta}$. The estimation consists in maximizing the simulated log-likelihood of the model. To simplify notation, let $\mathbf{z}_{ch} = (\mathbf{x}_c, \mathbf{y}_h)$ denote the observable characteristics of placement $(c, h) \in C_i \times H_i$, and group all the observable characteristics of market i in $\mathbf{Z}_i = (\mathbf{X}_i, \mathbf{Y}_i)$. Similarly, let $\mathbf{\Omega}_i = (\boldsymbol{\omega}_{ch})_{(c,h)\in C_i \times H_i}$.

Fix $\boldsymbol{\theta} \in \mathbf{\Theta}$. Consider an arbitrary market i. The likelihood of observing $(M_i, \mathbf{T}_i, \mathbf{R}_i)$, conditional on $(\mathbf{\Omega}_i, \mathbf{Z}_i)$, is given by:

$$\mathcal{L}(M_i, \mathbf{T}_i, \mathbf{R}_i \,|\, \mathbf{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) = \mathcal{L}_M(M_i \,|\, \mathbf{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_M, \boldsymbol{\theta}_T) \mathcal{L}_{\mathbf{T},\mathbf{R}}(\mathbf{T}_i, \mathbf{R}_i \,|\, M_i, \mathbf{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T) \tag{23}$$

where $\mathcal{L}_M(M_i \mid \mathbf{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_M, \boldsymbol{\theta}_T)$ denotes the conditional matching likelihood, and $\mathcal{L}_{\mathbf{T},\mathbf{R}}(\mathbf{T}_i, \mathbf{R}_i \,|\, M_i, \mathbf{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T)$ denotes the conditional outcome likelihood. Both likelihood functions are conditional on both unobservable and observable characteristics, $\mathbf{\Omega}_i$ and $\mathbf{Z}_i$, respectively. In the next two sections, I spell out both conditional likelihood functions. Then, I show how to compute the simulated log-likelihood of the data, which basically amounts to integrating out $\mathbf{\Omega}_i$ from (23).

5.1 Conditional Matching Likelihood

Write the payoff function $\pi(\cdot)$ as a function of placement characteristics and parameters, i..e, $\pi(c,h) = \pi(\boldsymbol{\omega}_{ch}, \mathbf{z}_{ch} \,|\, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M)$. Also, let $\mathbb{M}_i \equiv \mathbb{M}(C_i, H_i)$ denote the set of feasible matchings in market i. The conditional matching likelihood is given by the Probit choice probability:

$$\text{(24)} \qquad \mathcal{L}_M(M_i \,|\, \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) = \int 1_{\mathcal{A}(M_i|\boldsymbol{\Omega}_i,\mathbf{Z}_i,\boldsymbol{\theta}_T,\boldsymbol{\theta}_M)}(\boldsymbol{\upsilon}) dF(\boldsymbol{\upsilon}),$$

where $\boldsymbol{\upsilon} = (\upsilon_M)_{M\in\mathbb{M}_i}$ is the vector of matching composite errors, $1_{\mathcal{A}}(\boldsymbol{\upsilon})$ denotes the indicator function of set $\mathcal{A}$ (it takes $\boldsymbol{\upsilon}$ as argument), and the set $\mathcal{A}(M_i \mid \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M)$ is the set of $\boldsymbol{\upsilon}$'s for which the matching M_i is optimal, i.e.,

$$\text{(25)} \qquad \left\{ \boldsymbol{\upsilon} : \upsilon_M - \upsilon_{M_i} \leq \sum_{c,h} [M_i(c,h) - M(c,h)]\, \pi(\boldsymbol{\omega}_{ch}, \mathbf{z}_{ch} \,|\, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) \, \forall M \in \mathbb{M}_i \right\}.$$

5.2 Conditional Outcomes Likelihood

Let $\mathcal{L}_{T,R}(T, R \mid \boldsymbol{\omega}, \mathbf{z}, \boldsymbol{\theta}_T)$ denote the conditional likelihood of a single placement outcome, given by the Burr competing risks likelihood:

$$\text{(26)} \qquad \mathcal{L}_{T,R}(T, R \,|\, \boldsymbol{\omega}, \mathbf{z}, \boldsymbol{\theta}_T) = \bar{F}(T \,|\, \boldsymbol{\omega}, \mathbf{z}, \boldsymbol{\theta}_T) \lambda_R(T \,|\, \boldsymbol{\omega}, \mathbf{z}, \boldsymbol{\theta}_T)^{1_{R \notin \{em, cen\}}}$$

where $\bar{F}(T \,|\, \boldsymbol{\omega}, \mathbf{z}, \boldsymbol{\theta}_T)$ is the survivor function given in (18), and $\lambda_R(T \,|\, \boldsymbol{\omega}, \mathbf{z}, \boldsymbol{\theta}_T)$ the termination specific hazard-rate in Assumption 2. The conditional outcome likelihood of all the placements in market i is given by:

$$\text{(27)} \qquad \mathcal{L}_{\mathbf{T},\mathbf{R}}(\mathbf{T}_i, \mathbf{R}_i \,|\, M_i, \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T) = \prod_{(c,h)\in M_i} \mathcal{L}_{T,R}(T_{ch}, R_{ch} \,|\, \boldsymbol{\omega}_{ch}, \mathbf{z}_{ch}, \boldsymbol{\theta}_T).$$

5.3 (Simulated) Log-likelihood

Let $\mathcal{G}$ denote the joint distribution of $\boldsymbol{\Omega}_i$, i.e., $\mathcal{G} = \times_{c,h} G_{ch}$, where $G_{ch} \equiv N(0, \boldsymbol{\Sigma}_\omega)$. The conditional likelihood of the market-level data $(M_i, \mathbf{T}_i, \mathbf{R}_i)$ is:

$$\text{(28)} \qquad \mathcal{L}(M_i, \mathbf{T}_i, \mathbf{R}_i \,|\, \mathbf{Z}_i, \boldsymbol{\theta}) = \int \mathcal{L}_M(M_i \,|\, \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) \mathcal{L}_{\mathbf{T},\mathbf{R}}(\mathbf{T}_i, \mathbf{R}_i \,|\, M_i, \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T) \mathcal{G}(d\boldsymbol{\Omega}_i \,|\, \boldsymbol{\Sigma}_\omega).$$

The log-likelihood of the data is

$$\ell_n(\boldsymbol{\theta}\,|\,\mathbf{Z}) = \sum_{i=1}^{n} \log \mathcal{L}(M_i, \mathbf{T}_i, \mathbf{R}_i\,|\,\mathbf{Z}_i, \boldsymbol{\theta}). \tag{29}$$

To estimate $\boldsymbol{\theta}$, I compute the simulated counterpart of $\ell_n(\boldsymbol{\theta}\,|\,\mathbf{Z})$. There are two multi-dimensional integrals within (29) that need to be simulated. The first one is the integral over $\boldsymbol{\upsilon}$ in the conditional matching likelihood, see (24). To compute this integral, I draw a sample of S_υ independent draws of the taste variation terms, ε_c and η_h. The sample is drawn independently of the model parameters, in order to keep the simulation draws fixed during the estimation. I use a logit-kernel to smooth the choice probabilities in (24). It is well known (e.g. Train 2009), that such smoothing is computationally convenient when estimating multinomial probit models, especially in cases with a large number of alternatives, as in this case. Let $\zeta > 0$ denote the smoothing parameter of the logit-kernel. The second integral that needs to be computed through simulation is the one over $\boldsymbol{\Omega}_i$ in (28). To compute this integral, I draw a random sample of S_ω independent draws of each $\boldsymbol{\omega}_{ch} = (\omega_{R,ch})_{R\in\mathcal{R}_0}$, for $(c,h) \in C_i \times H_i$, $i = 1,\ldots,n$. Likewise, this sample is drawn independently of the model parameters. Let $\ell_n^{S_\omega,S_\upsilon,\zeta}(\boldsymbol{\theta}\mid\mathbf{Z})$ denote the simulated counterpart of the log-likelihood of the data in (29). (See Appendix A for more details on the estimation.) The estimator of $\boldsymbol{\theta}$ is given by:

$$\hat{\boldsymbol{\theta}}_{SMLE} = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \ell_n^{S_\omega,S_\upsilon,\zeta}(\boldsymbol{\theta}\,|\,\mathbf{Z}) \tag{30}$$

Standard results (e.g. Gourieroux and Monfort 1997) imply that $\hat{\boldsymbol{\theta}}_{SMLE}$ is a consistent, efficient, and asymptotically normal estimator for $\boldsymbol{\theta}$, as $n, S_\omega, S_\upsilon \to \infty$ with $\min\{S_\omega, S_\upsilon\}/\sqrt{n} \to \infty$.

6 Estimation Results

6.1 Empirical Specification

In this section I present the results of the estimation. Due to computational considerations, I consider a small version of the model in terms of the number of covariates I include. The estimates presented below, correspond to a model that includes the following placement characteristics: age, type of foster home (relative, county, agency, or group home), and distance to school. I also include a dummy for children for which the school's zip-code is missing (who presumably do not go to school), and interactions between age and the type of foster home.

I define children and home-types (used to specify the taste variation terms) as follows. The set of child-types, X, contains two elements differentiating children who are younger, or older, than 8 years old. The set of home-types, Y, includes one type for each type of foster home other than relatives. It is not necessary to define a home-type for relative foster homes since all of them are in singleton markets.

The dataset used in the estimation contains 1,467 markets, and 2,358 assigned placements. This specification of the model has 39 parameters.

6.2 Parameter Estimates

In this section, I discuss the simulated maximum likelihood estimates of the model parameters. Table 2 presents the parameter estimates of the outcome distribution, $\hat{\mathbf{\Sigma}}_\omega$ and $\hat{\boldsymbol{\theta}}_T$. The first two rows of the table present the estimated covariance matrix of $\boldsymbol{\omega}$. The estimated variance of ω_d is higher than that of ω_{ex}, implying that the variance not captured by the covariates is higher for disruption than for exit. The model also captures a positive correlation between both hazard rates: placements which the matchmaker considers as having a higher hazard for disruption, are also considered as having a slightly higher hazard for exiting the system.

The next rows of of Table 2 report the estimated coefficients of each of the covariates in $g(\mathbf{x}, \mathbf{y})$ for each hazard rate. A larger coefficient of (say) age on the disruption hazard implies that placements with older children are more likely to be disrupted (and sooner) than placements with younger children. The coefficients indicate that older children have higher disruption hazards in all types of foster homes, other than group homes. By contrast, age is found to have a minor effect in the hazard for exiting to permanency in foster homes other than group homes.

Table 3 reports average partial effects of placement characteristics on placement outcomes. Partial effects are computed for every placement assigned in the data using expressions (16) and (17). Here, one can see that, on average, placements with older children are more likely to be disrupted. The marginal effect of one year of

Table 2: Estimated Parameters of Outcome Distribution ($\mathbf{\Sigma}_\omega, \boldsymbol{\theta}_T$)

	(1) Disruption	(2) Exit
Var(ω_R)	0.873***	0.02955
	(0.2912)	(0.02867)
Cov(ω_d, ω_{ex})	0.1573*	0.1573*
	(0.08908)	(0.08908)
Age At Plac.	0.09872***	-0.01615
	(0.01767)	(0.01047)
County-FH	2.217***	-0.02375
	(0.332)	(0.2101)
Agency-FH	2.983***	0.4547***
	(0.2556)	(0.1237)
Group Home	-2.077**	-1.987***
	(0.9188)	(0.5642)
Age At Plac. × *County-FH*	-0.02272	0.01804
	(0.0261)	(0.01636)
Age At Plac. × *Agency-FH*	-0.07878***	-0.01007
	(0.0194)	(0.0124)
Age At Plac. × *Group Home*	0.2569***	0.1419***
	(0.06179)	(0.03894)
Distance To School (zip)	0.02052***	-0.006059***
	(0.002471)	(0.001724)
Missing Dist. To School	0.9007***	0.1222
	(0.1603)	(0.08942)
Constant	-8.996***	-6.082***
	(0.5408)	(0.2132)
Alpha (α_R)	1.091***	0.9665***
	(0.07551)	(0.03427)
Gamma (γ_R)	0.9527***	0.2222
	(0.1183)	(0.2361)
Number of markets (n)	1467	
$SMLL$	-17005.86	

Note: Estimated parameters of unobserved heterogeneity ($\mathbf{\Sigma}_\omega$) and conditional hazard rates ($\boldsymbol{\theta}_T$). Standard errors in parenthesis. Significance level of parameters: ***p<0.01, **p<0.05, *p<0.01.

Table 3: Average Partial Effects (APEs)

	(1) $\mathbb{P}$(Disrup)	(2) $\mathbb{P}$(Exit)	(3) $\mathbb{P}$(Emanc)	(4) $\mathbb{E}(\log T \mid \text{Disrup})$	(5) $\mathbb{E}(\log T \mid \text{Exit})$	(6) $\mathbb{E}(\log T)$
Age At Plac.	0.01393	-0.01146	-0.002465	-0.04059	-0.0218	-0.04014
County-FH	0.3168	-0.2661	-0.05067	-0.9689	-0.6275	-0.9266
Agency-FH	0.32	-0.2716	-0.04837	-1.221	-0.8743	-1.174
Group Home	0.1652	-0.1575	-0.007732	0.2872	0.4496	0.3393
Distance To School (zip)	0.004013	-0.003757	-0.0002561	-0.007978	-0.003091	-0.007359
Missing Dist. To School	0.1136	-0.09686	-0.0167	-0.5244	-0.3653	-0.5212
Number of placements				2358		

Note: Average partial effects of placement characteristics on expected outcomes. Averages taken across the sample of assigned placements in the data. The partial effects with respect to continuous variables is taken by considering a marginal change of one unit.

age on the disruption probability is, on average, 1.4%. Also, placements with older children are more likely to be disrupted sooner when they do so. Indeed, placements with older children tend to have lower durations overall, regardless of the termination reason. Placements with relatives are the more stable, they have lower disruption probabilities than every other type of foster home. They also last less than every other type of placement except for group homes. Placements in county and agency foster homes have similar expected outcomes. Both of them are around 30% more likely to be disrupted than placements with relatives. The distance between a foster home and the child's school increases the odds of disruption and overall diminishes a placement's expected duration.

Table 4 reports goodness of fit measures and the parameters used in the estimation. Overall, the model does good job on matching the average outcomes observed in the data. Note that when computing average expected outcomes for goodness of fit purposes, one must take into account censored placements (those for which the outcome is not observable due to the sample period). This is done by replacing T_{em} for $\min\{T_{em}, T_{cen}\}$ in expressions (16) and (17). Also, note that the (average) expected log-duration conditional on emancipation/censoring predicted by the model is much higher than the emancipation/censoring times observed in the data. This is reflects that the placements that are more likely to be emancipated or censored are precisely the ones that have lower times to emancipation or are closer to the end of the sample period.

Table 5 reports the estimated parameters of the matchmaker's utility function. Overall, the matchmaker has a higher payoff from placements that exit to permanency. The least desirable termination reason is disruption. The marginal utility of

Table 4: Goodness of Fit and Estimation Parameters

	(1) Predicted	(2) Sample
$\mathbb{P}$*(Disruption)*	0.514	0.5093
$\mathbb{P}$*(Exit)*	0.4303	0.4237
$\mathbb{P}$*(Emanc/Cens)*	0.05568	0.06701
$\mathbb{E}(\log T \mid \textit{Disruption})$	4.482	4.141
$\mathbb{E}(\log T \mid \textit{Exit})$	4.721	4.994
$\mathbb{E}(\log T \mid \textit{Emanc/Cens})$	7.19	5.534
$\mathbb{E}(\log T)$	4.615	4.596
Number of markets (n)	1467	
Number of assigned placements	2358	
Number of prospective placements	8900	
$SMLL$	-17005.86	
S_ω	50	
S_υ	50	
ζ	1e-01	
dim($\boldsymbol{\theta}$)	39	

Note: Average predicted outcomes and sample average outcomes. Averages taken across the sample of assigned placements in the data. The number of assigned placements in the data is equal to $\sum_i \sum_{c,h} M_i(c,h)$. The number of prospective placements is equal to $\sum_i \sum_{c,h} |C_i| \times |H_i|$. $SMLL$ gives the value of the simulated log-likelihood at the estimated vector of parameters. S_ω, S_υ, and ψ are the parameters of the simulated log-lilkelihood. $\dim(\boldsymbol{\theta})$ refers to the number of parameters estimated.

duration is negative, regardless of termination reason. The magnitude of the parameters show that the matchmaker is *not* willing to trade-off a placement exiting to permanency for it being disrupted, regardless of the time to reach permanency and the time spent in a disruptive placement. To see this, note that if a placement is to be disrupted, the matchmaker prefers for it to be disrupted as soon as possible. However, even if a placement is disrupted right away, $T = 1$, the payoff to the matchmaker is lower than if the child exits to permanency, regardless of the time the child needs to wait before exiting.

The marginal utility of the time to emancipation is positive conditional on duration, but negative conditional on exiting to permanency. This captures that the valuation of the matchmaker for age differs depending on the termination reason. An interpretation of this preference is that the time to disruption and permanency (i.e., the time that it takes for a placement to be disrupted or exit to permanency, conditional on that being its termination reason) is valued differently depending on the age of children. The sign of the coefficients indicate that the matchmaker's preference against children spending time in placements that will be disrupted is stronger for younger children than for older ones. By contrast, the matchmaker's tolerance for children waiting to exit to permanency is higher for younger children than for older ones. The magnitude of the coefficients allow to compute the marginal rate of substitution between duration and age. For instance, consider a child of average age, 8.7 years old, who is in a placement known to be disrupted. And set the disruption time at its conditional average, 5.4 months (165 days). A placement that is know to be disrupted, but has a child who is younger by one year, generates a higher payoff for the matchmaker as long as its duration is less than 5.9 months (180 days), 9.31% more. If the placement is known to be terminated because the child will exit to permanency, the opposite obtains. Again, consider a placement with a child who is 8.7 years old and, who is known will exit to permanency in the average time, 10 months (304 days). A placement with a child who is also known will exit to permanency, but who is one year *older*, generates a higher payoff to the matchmaker, as long as the child exits to permanency in no more than 10.2 months (312 days), 2.6% more.

Table 6 reports the estimated covariance matrices of the taste variation terms. Overall, the estimates show no significance variance in the taste variation parameters. Intuitively, this reflect that, given the current specification, the expected out-

Table 5: Estimated Parameters of Matching Utility ($\boldsymbol{\theta}_M$)

	(1) Disruption	(2) Exit	(3) Emancipation
μ_R — *MgU. Term. Reason*	-2.908***	2.449**	-2.057***
	(0.6972)	(1.091)	(0.7183)
φ_R — *MgU. Duration*	-0.3549***	-0.5265***	0†
	(0.1005)	(0.167)	(0)
$\bar{\varphi}_R$ — *MgU. Emanc. Time*	0.3093***	-0.1179	0.009985
	(0.06172)	(0.09607)	(0.01364)
Number of markets (n)		1467	
$SMLL$		-17005.86	

Note: Estimated parameters of matching utility function ($\boldsymbol{\theta}_M$), where $u = \mu_R + \varphi_R \log T + \bar{\varphi}_R \log T_{em}$. Standard errors in parenthesis. Significance level of parameters: ***p<0.01, **p<0.05, *p<0.01. † indicates fixed parameter (i.e. not estimated).

Table 6: Estimates of the covariance matrix of the taste variation shocks ($\boldsymbol{\Sigma}$)

$$\hat{\boldsymbol{\Sigma}}_\varepsilon = \begin{pmatrix} 0\dagger & 0\dagger & 0\dagger \\ (0) & (0) & (0) \\ 0\dagger & 0.0002013 & -0.001219 \\ (0) & (0.0009768) & (0.003017) \\ 0\dagger & -0.001219 & 0.01181 \\ (0) & (0.003017) & (0.01172) \end{pmatrix}_{|Y|\times|Y|} \qquad \hat{\boldsymbol{\Sigma}}_\eta = \begin{pmatrix} 0\dagger & 0\dagger \\ (0) & (0) \\ 0\dagger & 0.0001188 \\ (0) & (0.000899) \end{pmatrix}_{|X|\times|X|}$$

Note: Estimated parameters of the covariance matrices of taste variation shock of children over home types, $\varepsilon_c = (\varepsilon_{cy})_{y\in Y} \sim N(0, \boldsymbol{\Sigma}_\epsilon)$, and of the covariance matrix of the taste variation shock of homes over children types, $\eta_h = (\eta_{xh})_{x\in X} \sim N(0, \boldsymbol{\Sigma}_\eta)$. Standard errors in parenthesis. Significance level of parameters: ***p<0.01, **p<0.05, *p<0.01. † indicates fixed parameter (i.e. not estimated).

comes of placements seem to be sufficient in order to predict placement assignments.[17]

7 Counterfactual Exercises

7.1 Counterfactual I — Market Thickness

In this section, I analyze the effect of policies aimed at improving outcomes by increasing market thickness. Market thickness may be increased along two dimen-

[17] A caveat of the estimates in Table 6 is that the normalizations implemented in this specification do not correspond to the ones in Assumption 4, which key in proving Proposition 1. The estimates in Table 6 may be close to zero because the normalizations are not doing a good job in identifying the parameters. Further estimations will implement the normalization specified in Assumption 4.

sions. First, I consider the case in which placements every $D > 1$ days, instead of daily as is done in the field ($D = 0$). I consider policies with $D \leq 15$. Second, I consider the case in which non-relative placements are assigned across all regional offices together, instead of within them as is done in the field. I also consider the two types of policies together, i.e., assigning placements every $D > 1$ days and pooling the children and foster homes from all regional offices into a county-wide market.

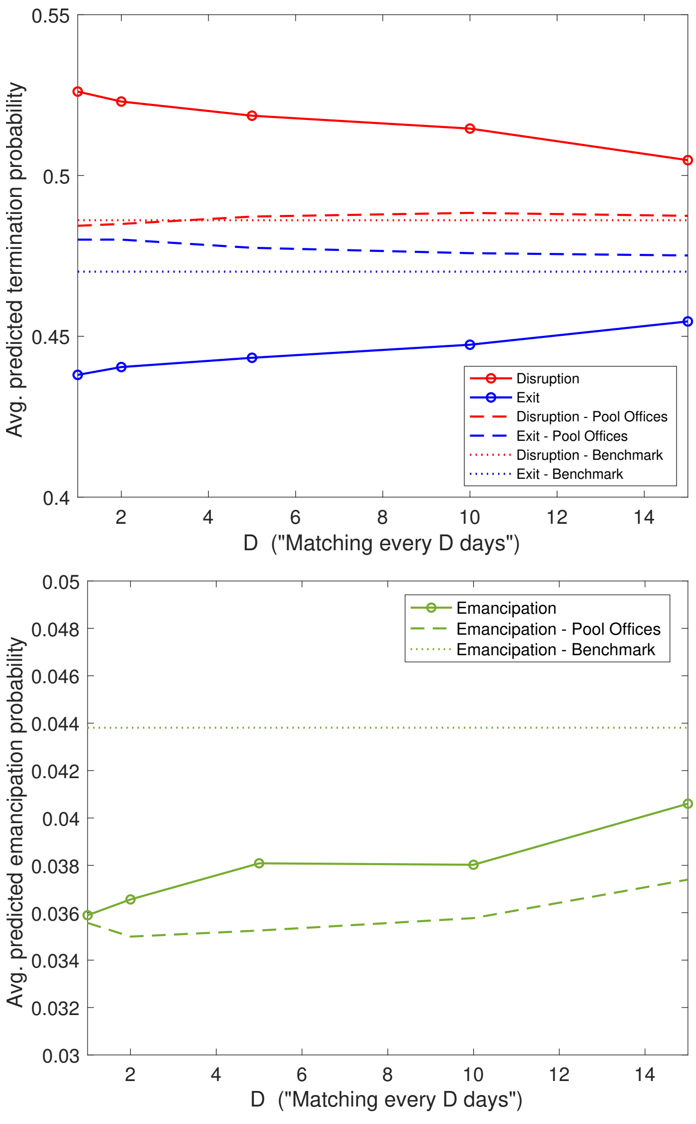
By design, the aggregate payoff of the matchmaker is higher when markets are thinker. The reason is because the original matching is always feasible when the market is thicker. The effect on the expected outcomes of placements is controlled by the matchmaker's payoff function, which determines which placements are assigned in the counterfactual markets.

Figure 1 plots the average predicted termination probabilities across the counterfactual markets. The value of D is plotted in the x-axis. The solid lines correspond to the termination probabilities in the case in which markets are formed within offices. The dashed lines to the case in which markets are pooled across regional offices. The plots also include a dotted line, which is constant across D. The dotted line corresponds to the "benchmark" case in which all placements are assigned at once, $D = \infty$, and regional offices are pooled together. The average predicted outcomes in the benchmark case correspond to the ones of the best placements (from the matchmaker's perspective) that can be formed in the full dataset.

The baseline values of the termination probabilities are the values at $D = 0$, i.e., these values correspond to the predicted probabilities of the model with the assignment observed in the data. From the top panel of Figure 1, one can see that in thicker markets the average disruption probability is lower, and that of exiting to permanency or disruption is higher. When the pools of available children and foster homes are larger, the matchmaker is able to assign placements with lower disruption probabilities. However, note that the gains from thicker markets come almost exclusively from pooling regional offices together. When matchings are assigned daily ($D = 0$) but regional offices are pooled together, the disruption rates diminishes from 52.61% to 48.43%. In terms of expectd number of placements per child, this is equivalent from going from 2.11 to 1.94.[18]

[18]The average disruption probability can be seen as the probability of a "failure" in a series of discrete dichotomic random draws. In which case, the number of placements per child follows a geometric distribution ("number of trials needed to get one success"). If p_d denotes the disruption probability, the expected number of placements per child is $1/(1 - p_d)$.

Figure 1: Counterfactual I — Average Predicted Termination Probabilities



Note: Plot of the average predicted termination termination probabilities. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of D, the number of non-matching days in a matching period. Solid lines correspond to counterfactuals in which markets are defined by regional offices; dashed lines to ones in which regional offices are pooled together into the same markets. The benchmark case (dotted line) corresponds to the case in which regional offices are pooled together into markets and $D = \infty$.

Figure 2 is analogous to Figure 1, but it plots the average predicted conditional durations of placements. Here, one can see that the average duration of placements may be higher or lower than the baseline in thicker markets. Interestingly, when offices are pooled together and placements are assigned daily, the matchmaker assigns placements with higher expected durations than both the baseline and benchmark cases. The reason is because the matchmaker is willing to trade-off duration (which it dislikes) with better termination probabilities. The same can be seen in Figure 3, which plots the average expected duration.

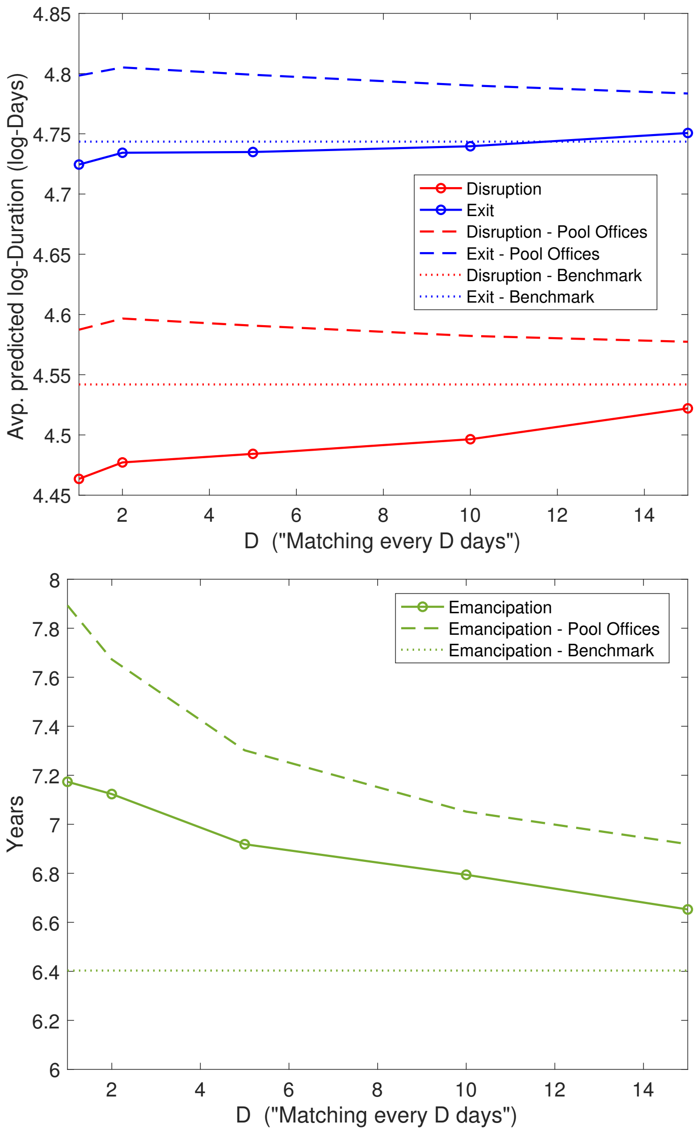
The top panel of Figure 4 plots the average distance to school across placements in thicker markets. The average distance between foster homes and children's schools is cut in 54% when offices are pooled together into county wide-markets. The average distance goes from 20.43 to 9.5 miles. From the plot, one can see that the gains resulting from lower disruption probabilities follows the same patters as the distance to school: the gains from pooling offices together outweighs those obtained from delaying placement assignments. Finally, the bottom plot of Figure 4 shows the average time that children waiting before being assigned placements. As expected, delaying placements increases this figure monotonically.

7.2 Counterfactual IIa — Relative Foster Homes

In this section, I analyze the effect that relatives have on average expected outcomes. Specifically, I consider an increase in the share of the foster homes that are relatives across all markets. I analyze both an increase in the intensive and extensive margins. Let $\delta_{rel} \in (0,1)$ be the increase in the share of foster homes that are relatives. I consider policies with $\delta_{rel} \leq 0.25$; $\delta_{rel} = 0$ corresponds to the baseline case, in which the supply of foster homes is the same as the one observed in the data.

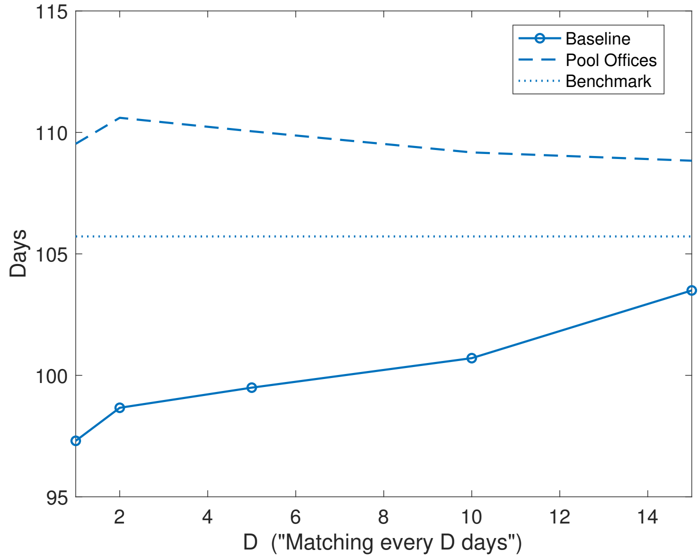
I increase the share of relative homes as follows. First, I estimate a binary logit model that predicts whether a child has a relative home, or not, as a function of its characteristics. Let $n^*_{rel} = \lfloor \delta_{rel} * n_{rel} \rfloor$, where n_{rel} denotes the number of placements with relatives in the data. Then, from the population of non-relative placements in the data, I select n^*_{rel} at random, weighting them by the predicted probability that each of them had a relative available. That is, I select children who did not had a relative placement, but had a higher likelihood of having it, with higher probability. In the case of the intensive margin, I convert the foster homes of the selected place-

Figure 2: Counterfactual I — Average Predicted Conditional Expected Duration



Note: Plots of the average predicted conditional expected durations. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of D, the number of non-matching days in a matching period. Solid lines correspond to counterfactuals in which markets are defined by regional offices; dashed lines to ones in which regional offices are pooled together into the same markets. The benchmark case (dotted line) corresponds to the case in which regional offices are pooled together into markets and $D = \infty$.

Figure 3: Counterfactual I — Average Predicted Expected Duration

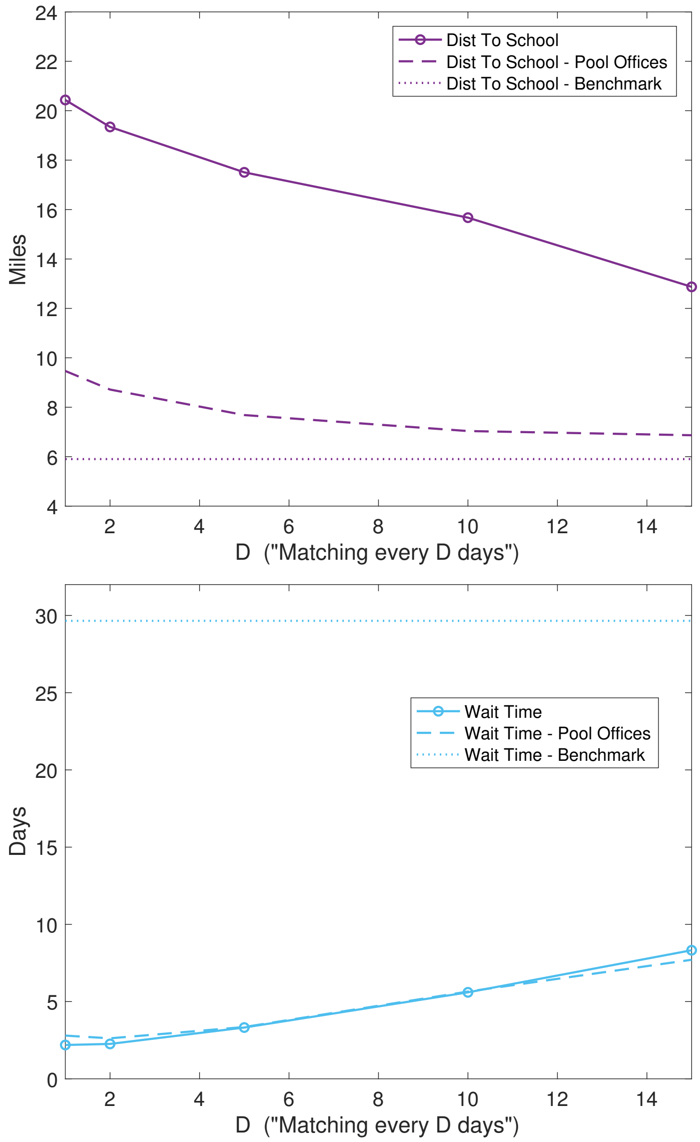


Note: Plot of the average predicted expected duration. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of D, the number of non-matching days in a matching period. Solid lines correspond to counterfactuals in which markets are defined by regional offices; dashed lines to ones in which regional offices are pooled together into the same markets. The benchmark case (dotted line) corresponds to the case in which regional offices are pooled together into markets and $D = \infty$.

ments into relative homes (leaving all other placement characteristics fixed), and assign them to new singleton markets with the corresponding child. In the extensive margin case, I create a duplicate of the foster homes of the selected children. Then, I convert the duplicated home to a relative home (leaving all the other placement characteristics fixed), and assign it with the corresponding child to a new singleton market. The difference between the intensive and extensive margins is that the set of available foster homes for the rest of the children in the market remains unchanged in the extensive margin, while it is reduced by one home in the intensive margin.

Figure 5 report the predicted average termination probabilities in the distinct counterfactuals. The parameter δ_{rel} is on the x-axis. One can observe that a higher share of relative homes, in both the intensive and extensive margins, has a sizable effect on termination probabilities. Overall, the disruption probability diminishes and the one of exiting to permanency increases. The adjustment is more gradual in the extensive margin. In the intensive margin, the disruption probability goes from 52.6% at $\delta_{rel} = 0$ to 45.82% at $\delta_{rel} = 0.25$ (equivalent to going from an average of 2.1 placement per child to 1.84). In the extensive margin, the change is from 52.6% to 47.8$ (equivalent to going from an average of 2.1 placement per child to 1.91).

Figure 4: Counterfactual I — Average Distance To School and Waiting Time



Note: Plots of the average distance to school and waiting time. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of D, the number of non-matching days in a matching period. Solid lines correspond to counterfactuals in which markets are defined by regional offices; dashed lines to ones in which regional offices are pooled together into the same markets. The benchmark case (dotted line) corresponds to the case in which regional offices are pooled together into markets and $D = \infty$.

The difference between both margins has to do in how the rest of the children are being placed in the non-relative placements. Figure 6 shows the analogous plots for conditional durations. Overall, placements tend to last longer when the share of relative foster homes in the system is increased.

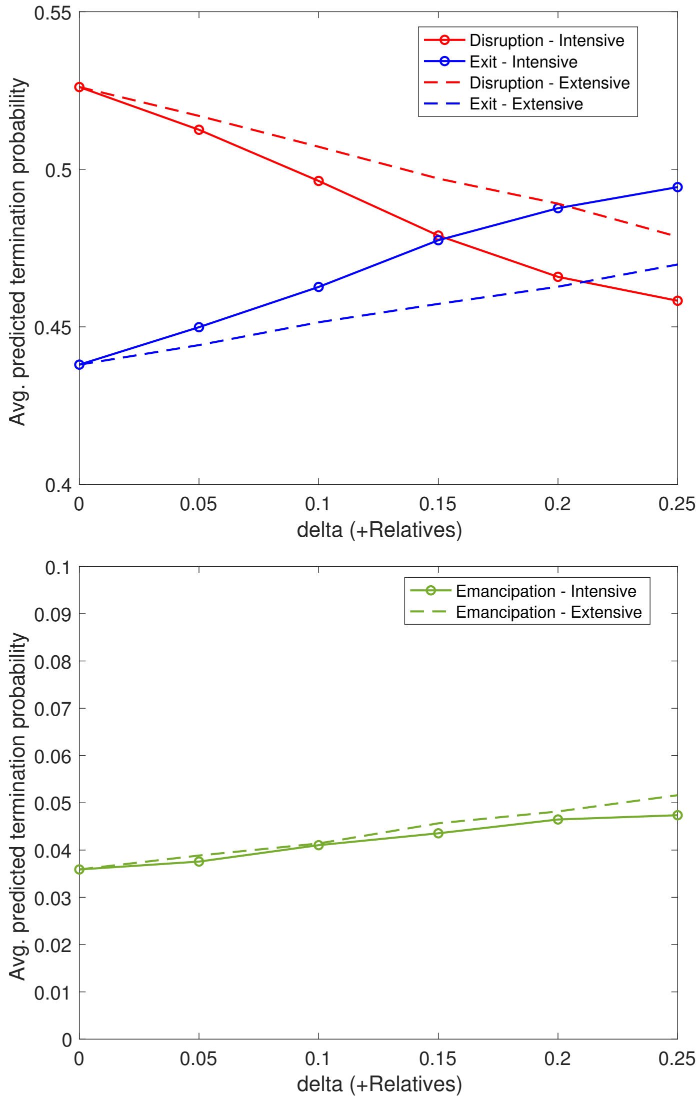
7.3 Counterfactual IIb — Agency Foster Homes

In this section, I analyze the effect that agency foster homes have on average expected outcomes. Specifically, I consider an increase in the share of foster homes that come through non-profit agencies across all markets. I analyze both an increase in the intensive and extensive margins. Let $\delta_{ah} \in (0, 1)$ be the increase in the share of foster homes that are agency homes. I consider policies with $\delta_{ah} \leq 0.25$; $\delta_{rel} = 0$ corresponds to the baseline case, in which the supply of foster homes is the same as the one observed in the data.

I increase the share of relative homes as follows. Let $n^*_{ah} = \lfloor \delta_{ah} * n_{ah} \rfloor$, where n_{ah} denotes the number of placements with agency homes in the data. Then, from the population of non-relative placements in the data, I select n^*_{ah} uniformly at random, keeping the relative share of the other types of non-relative placements fixed. In the case of the intensive margin, I convert the foster homes of the selected placements into relative homes (leaving all other placement characteristics fixed). In the extensive margin case, I create a duplicate of the foster homes of the selected children. Then, I convert the duplicated home to a relative home (leaving all the other placement characteristics fixed). The difference between the intensive and extensive margins is that the set of available foster homes in the market has the same number of homes in the intensive margin (with one converted into an agency home), and in the extensive margin it has an extra agency home.

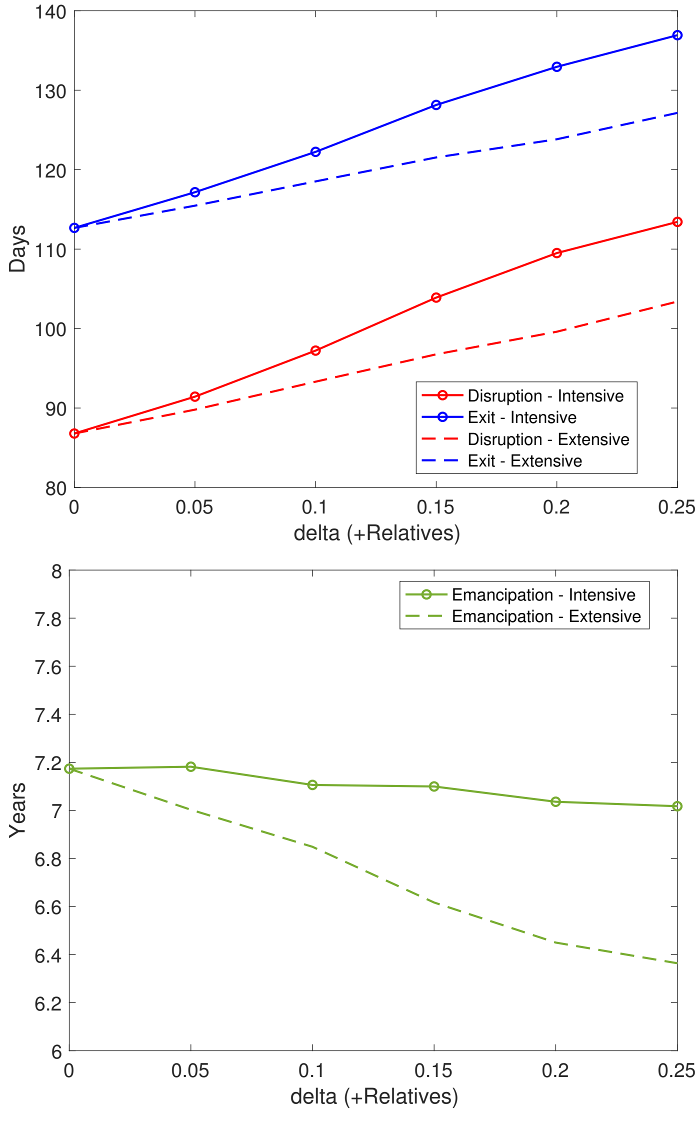
Figure 8 report the predicted average termination probabilities in the distinct counterfactuals. The parameter δ_{ah} is on the x-axis. One can observe that a higher share of agency homes, in both the intensive and extensive margins, has a minor effect on termination probabilities. Interestingly, the direction of the effects fo in opposite directions in the intensive and the extensive margins. In the intensive margin, the disruption probability diminishes and the one of exiting to permanency increases.The disruption probability goes from 52.6% at $\delta_{rel} = 0$ to 51% at $\delta_{rel} = 0.25$ (equivalent to going from an average of 2.1 placement per child to 2). In the ex-

Figure 5: Counterfactual IIa-Relatives — Average Predicted Termination Probabilities



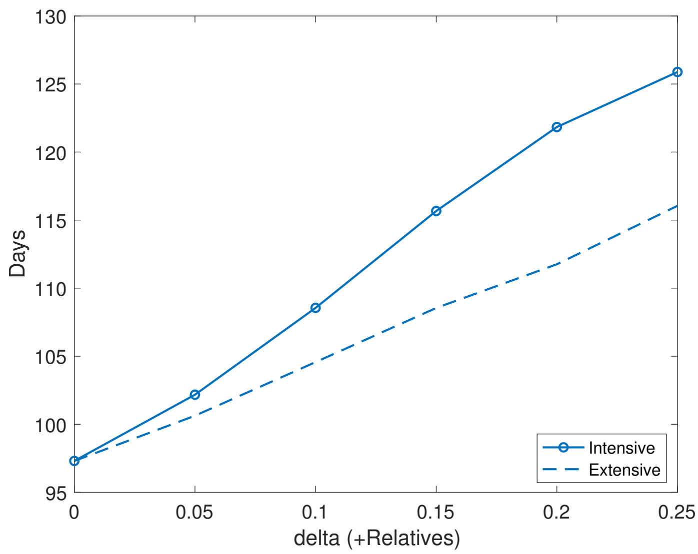
Note: Plot of the average predicted termination termination probabilities. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of δ_{rel}, the factor by which the supply of Relative Foster Homes is adjusted. Solid lines correspond to counterfactuals in which the supply of Relative Foster Homes is increased in the intensive margin; dashed lines to ones in which the supply of Relative Foster Homes is increased in the extensive margin.

Figure 6: Counterfactual IIa-Relatives — Average Predicted Conditional Expected Duration



Note: Plots of the average predicted conditional expected durations. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of δ, the factor by which the supply of Relative Foster Homes is adjusted. Solid lines correspond to counterfactuals in which the supply of Relative Foster Homes is increased in the intensive margin; dashed lines to ones in which the supply of Relative Foster Homes is increased in the extensive margin.

Figure 7: Counterfactual IIa-Relatives — Average Predicted Expected Duration



Note: Plot of the average predicted expected duration. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of δ, the factor by which the supply of Relative Foster Homes is adjusted. Solid lines correspond to counterfactuals in which the supply of Relative Foster Homes is increased in the intensive margin; dashed lines to ones in which the supply of Relative Foster Homes is increased in the extensive margin.

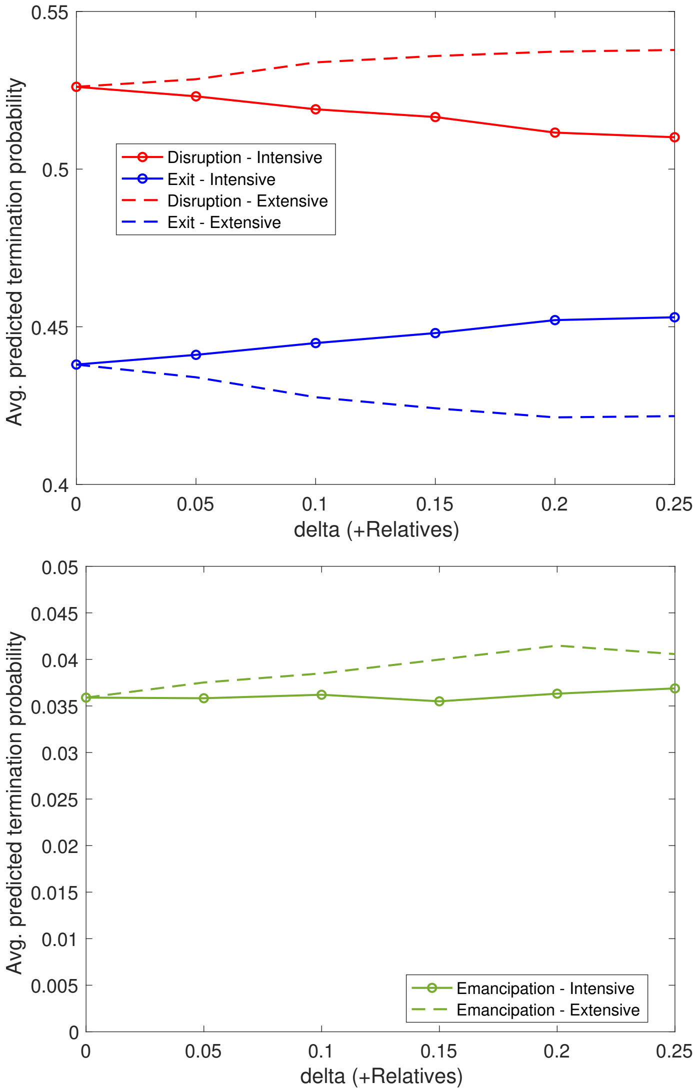
tensive margin, the average disruption probability increases, it goes from 52.6% to 53.8$ (equivalent to going from an average of 2.1 placement per child to 2.2). Figure 6 shows the analogous plots for conditional durations.

8 Conclusion

Often, the allocation of resources is the result of individual choices made within exogenously-designed institutions. This paper presents a framework to study how placements are assigned in foster care. The model aims to capture how social workers assign placements in the field. The model incorporates key institutional features of placement assignment in foster care: (1) children need to be placed with relatives whenever possible; (2) social workers need to prioritize the location of prospective foster homes in relation to the children's schools, and (3) social workers have discretion in how to weigh all the factors that contribute to successful placements.

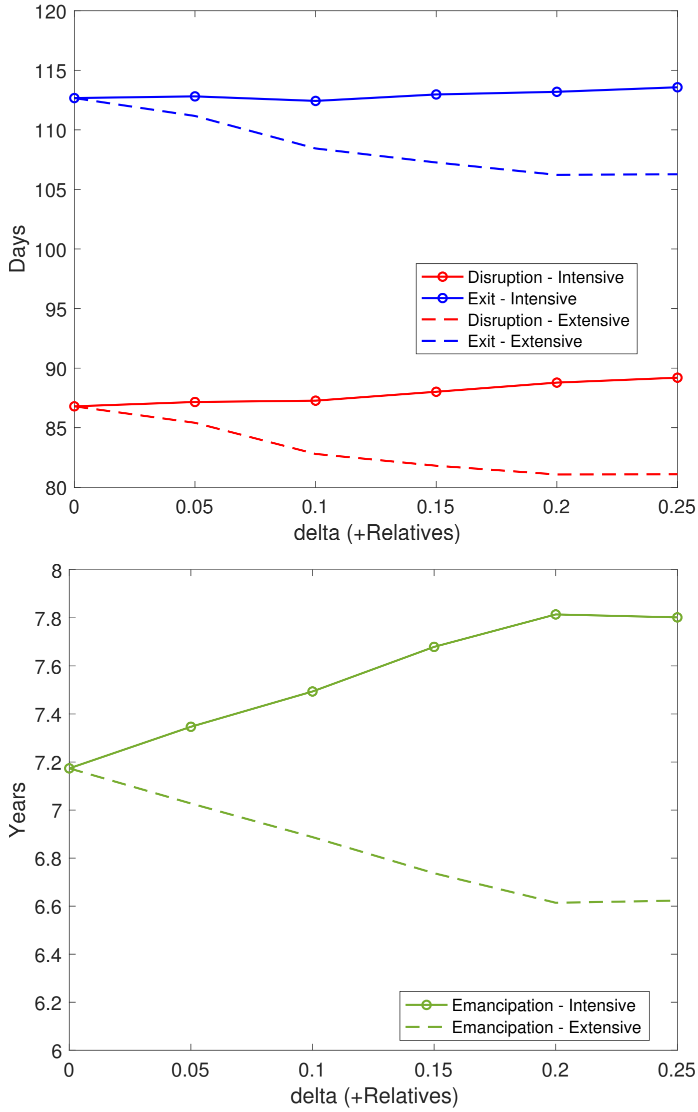
A key aspect of the model is that it incorporates the endogeneity arising from placement assignments being affected by unobservables correlated with outcomes. The main identification strategy of the paper is to rely on the exogenous variation

Figure 8: Counterfactual IIb-Agency-FH — Average Predicted Termination Probabilities



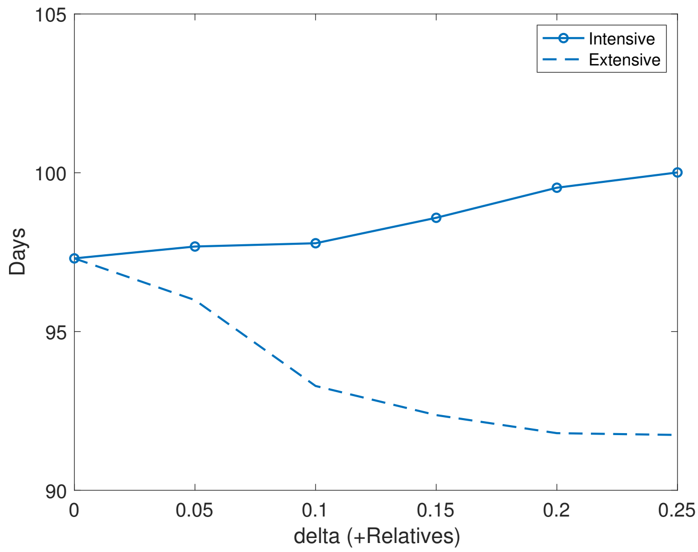
Note: Plot of the average predicted termination termination probabilities. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of δ, the factor by which the supply of Agency Foster Homes is adjusted. Solid lines correspond to counterfactuals in which the supply of Relative Foster Homes is increased in the intensive margin; dashed lines to ones in which the supply of Agency Foster Homes is increased in the extensive margin.

Figure 9: Counterfactual IIb-Agency-FH — Average Predicted Conditional Expected Duration



Note: Plots of the average predicted conditional expected durations. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of δ, the factor by which the supply of Relative Foster Homes is adjusted. Solid lines correspond to counterfactuals in which the supply of Agency Foster Homes is increased in the intensive margin; dashed lines to ones in which the supply of Agency Foster Homes is increased in the extensive margin.

Figure 10: Counterfactual IIb-Agency-FH — Average Predicted Expected Duration



Note: Plot of the average predicted expected duration. Averages taken across all assigned placements in each counterfactual. The x-axis plots the value of δ, the factor by which the supply of Agency Foster Homes is adjusted. Solid lines correspond to counterfactuals in which the supply of Agency Foster Homes is increased in the intensive margin; dashed lines to ones in which the supply of Relative Foster Homes is increased in the extensive margin.

across the dates and geographic regions at which children enter foster care. The empirical exercise uses a novel dataset of confidential foster care record from Los Angeles County, California. The parameter estimates of the model show that expected outcomes are significant factors when assigning placements. Overall, social workers assign the placements that are less likely to be disrupted, and in which it is more likely that the children exit to permanency. Another key variable when determining assignments is the conditional expected duration of prospective placements. Social workers aim to assign placements that, conditional on their termination reason, will have the lowest possible durations.

Through counterfactual exercises, I show the effect of market thickness and the presence of different types of foster homes on the distribution of outcomes. A key contribution of this paper is to quantify the gains, in terms of better placement outcomes, resulting from thicker markets if foster care. It is shown that the gains due to market thickness are greater when thickness is increased geographically (by assigning placements throughout the county) than time-wise (by delaying placements). Specifically, the model predicts that if placements were assigned in county-wide markets, the expected number of placements children would experience in foster care would diminish by 8%, and the average distance from foster homes to children's schools would be reduces by 54%.

As a final note, it is worthwhile emphasizing how the findings of this paper supports the view that social workers have a good understanding of which placements are less likely to be disrupted, and seem to do a pretty good job when it comes to assigning them. They assign the placements that are more likely work. However, at the system level, the model shows that the current state of the system does not facilitate the coordination between the distinct regional offices. The paper finds evidence that by being better at coordinating with one another, regional offices would be able to assign better placements for children and foster parents.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A. Pathak. 2017. "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match." *American Economic Review* 107, no. 12 (December): 3635–89. (Cited on page 6).

Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez. 2005. "The Boston Public School Match." *American Economic Review* 95 (2): 368–371. (Cited on page 6).

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review* 93, no. 3 (June): 729–747. (Cited on page 6).

Ackerberg, Daniel A., and Maristella Botticini. 2002. "Endogenous Matching and the Empirical Determinants of Contract Form." *Journal of Political Economy* 110 (3): 564–591. (Cited on pages 23, 24).

Agarwal, Nikhil. 2015. "An Empirical Model of the Medical Match." *American Economic Review* 105 (7): 1939–78. (Cited on page 5).

Agarwal, Nikhil, Itai Ashlagi, Eduardo Azevedo, Clayton Featherstone, and Ömer Karaduman. 2017. "Market Failure in Kidney Exchange." Unpublished paper. (Cited on page 6).

Agarwal, Nikhil, Itai Ashlagi, Michael Rees, Paulo Somaini, and Daniel Waldinger. 2019. "An Empirical Framework for Sequential Assignment: The Allocation of Deceased Donor Kidneys." Unpublished paper. (Cited on page 6).

Agarwal, Nikhil, and Paulo Somaini. 2018. "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism." *Econometrica* 86 (2): 391–444. (Cited on page 5).

Akbarpour, Mohammad, Shengwu Li, and Shayan Oveis Gharan. Forthcoming. "Thickness and Information in Dynamic Matching Markets." *Journal of Political Economy.* (Cited on page 6).

Ashlagi, Itai, Patrick Jaillet, and Vahideh H. Manshadi. 2013. "Kidney Exchange in Dynamic Sparse Heterogenous Pools." Unpublished paper. (Cited on page 6).

Baccara, Mariagiovanna, Allan Collard-Wexler, Leonardo Felli, and Leeat Yariv. 2014. "Child-Adoption Matching: Preferences for Gender and Race." *American Economic Journal: Applied Economics* 6, no. 3 (July): 133–58. (Cited on page 5).

Baccara, Mariagiovanna, SangMok Lee, and Leeat Yariv. Forthcoming. "Optimal Dynamic Matching." *Theoretical Economics.* (Cited on page 6).

Buchholz, Nicholas. 2019. "Spatial Equilibrium, Search Frictions and Dynamic Efficiency in the Taxi Industry." Unpublished paper. (Cited on page 5).

Calsamiglia, Caterina, Chao Fu, and Maia Güell. 2017. "Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives." Unpublished paper. (Cited on page 6).

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2012. "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market." *Journal of Political Economy* 120 (4): 659–695. (Cited on page 5).

Choo, Eugene. 2015. "Dynamic Marriage Matching: An Empirical Framework." *Econometrica* 83 (4): 1373–1423. (Cited on page 5).

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy* 114 (1): 175–201. (Cited on page 5).

DCFS. 2019. *Los Angeles County Department of Children and Family Services Child Welfare Policy Manual.* URL: http://policy.dcfs.lacounty.gov. (Cited on pages 7, 8).

Doval, Laura. 2018. "A Theory of Stability in Dynamic Matching Markets." Unpublished paper. (Cited on page 6).

Doyle Jr., Joseph J. 2007. "Child Protection and Child Outcomes: Measuring the Effects of Foster Care." *American Economic Review* 97, no. 5 (December): 1583–1610. (Cited on page 5).

———. 2008. "Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care." *Journal of Political Economy* 116 (4): 746–770. (Cited on page 5).

———. 2013. "Causal effects of foster care: An instrumental-variables approach." *Children and Youth Services Review* 35 (7): 1143–1151. (Cited on page 5).

Doyle Jr., Joseph J., and Anna Aizer. 2018. "Economics of Child Protection: Maltreatment, Foster Care, and Intimate Partner Violence." *Annual Review of Economics* 10 (1): 87–108. (Cited on page 5).

Doyle Jr., Joseph J., and H. Elizabeth Peters. 2007. "The market for foster care: an empirical study of the impact of foster care subsidies." *Review of Economics of the Household* 5 (4): 329–351. (Cited on page 5).

Dregan, Alexandru, and Martin C. Gulliford. 2012. "Foster care, residential care and public care placement patterns are associated with adult life trajectories: population-based cohort study." *Social Psychiatry and Psychiatric Epidemiology* 47 (9): 1517–1526. (Cited on page 3).

Ewens, Michael, Alex Gorbenko, and Arthur Korteweg. 2019. "Venture Capital Contracts." Unpublished paper. (Cited on page 24).

FAM. 2019. *California — Family Code.* URL: `http://leginfo.legislature.ca.gov/faces/codesTOCSelected.xhtml?tocCode=FAM&tocTitle=+Family+Code+-+FAM`. (Cited on pages 7, 8).

Fox, Jeremy T. 2016. "Estimating Matching Games with Transfers." Unpublished paper. (Cited on page 5).

Fréchette, Guillaume R., Alessandro Lizzeri, and Tobias Salz. 2019. "Frictions in a Competitive, Regulated Market: Evidence from Taxis." *American Economic Review* 109, no. 8 (August): 2954–92. (Cited on page 5).

Gale, David, and Lloyd S. Shapley. 1962. "College Admissions and the Stability of Marriage." *The American Mathematical Monthly* 69 (1): 9–15. (Cited on page 6).

Galichon, Alfred, and Yu-Wei Hsieh. 2017. "A Theory of Decentralized Matching Markets Without Transfers, with an Application to Surge Pricing." Unpublished paper. (Cited on page 5).

Galichon, Alfred, and Bernard Salanié. 2015. "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models." Unpublished paper. (Cited on page 5).

Gourieroux, Christian, and Alain Monfort. 1997. *Simulation-based Econometric Methods.* CORE Lectures. Oxford: Oxford University Press. (Cited on page 27).

Graham, Bryan S. 2011. "Chapter 19 - Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers." In *Handbook of Social Economics, vol. 1,* edited by Gharan Benhabib, Alberto Bisin, and Matthew O. Jackson, 965–1052. North-Holland. (Cited on page 5).

———. 2013. "Comparative Static and Computational Methods for an Empirical One-to-one Transferable Utility Matching Model." In *Structural Econometric Models,* edited by Eugene Choo and Matthew Shum, 153–181. Emerald Group Publishing Limited. (Cited on page 5).

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153–161. (Cited on page 22).

Heckman, James J., and Bo E. Honoré. 1989. "The Identifiability of the Competing Risks Model." *Biometrika* 76 (2): 325–330. (Cited on page 22).

Hwang, Sam I. M. 2016. "A Robust Redesign of High School Match." Unpublished paper. (Cited on page 6).

Kalbfleisch, John D., and Ross L. Prentice. 2002. *The Statistical Analysis of Failure Time Data, Second Edition.* New Jersey: John Wiley & Sons, Inc. (Cited on page 20).

Lancaster, Tony. 1990. *The Econometric Analysis of Transition Data.* Cambridge: Cambridge University Press. (Cited on pages 16, 20, 23).

Narita, Yusuke. 2016. "Match or Mismatch: Learning and Inertia in School Choice." Unpublished paper. (Cited on page 6).

NDACAN. 2015. "Adoption and Foster Care Analysis & Report System (AFCARS) Database Foster Care File FY 2015v1." *National Data Archive on Child Abuse and Neglect (NDACAN).* (Cited on pages 3, 7).

Roth, Alvin E. 1984. "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory." *Journal of Political Economy* 92, no. 6 (December): 991–1016. (Cited on page 6).

———. 2016. *Who Gets What — and Why.* New York: First Mariner Books. (Cited on page 5).

Rymph, Catherine E. 2017. *Raising Government Children. A History of Foster Care and the American Welfare State.* Chapel Hill: The University of North Carolina Press. (Cited on page 7).

Shapley, Lloyd S., and Martin Shubik. 1971. "The assignment game I: The core." *International Journal of Game Theory* 1 (1): 111–130. (Cited on page 5).

Singh, S. K., and G. S. Maddala. 1976. "A Function for Size Distribution of Incomes." *Econometrica* 44 (5): 963–970. (Cited on page 16).

Slaugh, Vincent, Mustafa Akan, Onur Kesten, and Utku M. Ünver. 2015. *The Pennsylvania Adoption Exchange Improves Its Matching Process.* Boston College Working Papers in Economics 858. Boston College Department of Economics. (Cited on page 4).

Sørensen, Morten. 2007. "How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital." *The Journal of Finance* 62 (6): 2725–2762. (Cited on page 24).

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press. (Cited on pages 24, 27, 55).

UC Davis. 2008. *A Literature Review of Placement Stability in Child Welfare Service: Issues, Concerns, Outcomes and Future Directions.* The University of California, Davis, Extension, The Center for Human Services. (Cited on page 3).

Ünver, M. Utku. 2010. "Dynamic Kidney Exchange." *The Review of Economic Studies* 77 (1): 372–414. (Cited on page 6).

Waldinger, Daniel. 2019. "Targeting In-Kind Transfers Through Market Design: A Revealed Preference Analysis of Public Housing Allocation." Unpublished paper. (Cited on page 6).

WIC. 2019. *California — Welfare and Institutions Code.* URL: http://leginfo.legislature.ca.gov/faces/codesTOCSelected.xhtml?tocCode=WIC&tocTitle=+Welfare+and+Institutions+Code+-+WIC. (Cited on pages 7, 8).

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data.* Cambridge: MIT Press. (Cited on page 16).

A Appendix: Estimation Details

In this section, I show the steps to compute the simulated log-likelihood $\ell_n^{S_\omega,S_\upsilon,\psi}(\boldsymbol{\theta} \,|\, \mathbf{Z})$ in detail, see (30).

1. Simulate the conditional matching likelihood, using a logit-kernel:

   - Write the surplus of matching $M \in \mathbb{M}_i$ as:

   $$V_{s_\upsilon}(M \mid \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) = \sum_{c,h} M(c,h) \left[\pi(\boldsymbol{\omega}_{ch}, \mathbf{z}_{ch} \,|\, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) + \varepsilon^{s_\upsilon}_{cy_h} + \eta^{s_\upsilon}_{x_c h}\right],$$

   where $\varepsilon_c^{s_\upsilon} = \left(\varepsilon_{cy}^{s_\upsilon}\right)_{y \in Y}$ and $\eta_h^{s_\upsilon} = \left(\eta_{xh}^{s_\upsilon}\right)_{x \in X}$ are simulated structural errors.

   - To simulate $\varepsilon_c^{s_\upsilon}$ and $\eta_h^{s_\upsilon}$, let $\boldsymbol{\Gamma}_\varepsilon$ and $\boldsymbol{\Gamma}_\eta$ be the Cholesky factors of $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$, respectively. Draw fixed simulated values

   $$\begin{aligned} \tilde{\varepsilon}_c^{s_\upsilon} &\sim \text{iid } N(0, I_{|Y|}), \quad s_\upsilon = 1, \ldots, S_\upsilon, \\ \tilde{\eta}_h^{s_\upsilon} &\sim \text{iid } N(0, I_{|X|}), \quad s_\upsilon = 1, \ldots, S_\upsilon, \end{aligned}$$

   for every $c \in C_i$ and $h \in H_i$. Set $\varepsilon_c^{s_\upsilon} = \boldsymbol{\Gamma}_\varepsilon \tilde{\varepsilon}_c^{s_\upsilon}$, and $\eta_h^{s_\upsilon} = \boldsymbol{\Gamma}_\eta \tilde{\eta}_h^{s_\upsilon}$.

   - Define the simulated counterpart of the conditional matching likelihood $\mathcal{L}_M(M_i \,|\, \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M)$ as

   $$\mathcal{L}_M^{s_\upsilon,\psi}(M_i \,|\, \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) = \frac{\exp\left\{V_{s_\upsilon}(M_i \mid \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M)/\psi\right\}}{\sum_{M \in \mathbb{M}_i} \exp\left\{V_{s_\upsilon}(M \mid \boldsymbol{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M)/\psi\right\}},$$

where $\psi > 0$ is the smoothing constant of the logit-kernel.

- Note that $\mathcal{L}_M^{s_\upsilon,\psi} = 1$ for all $M_i \in \mathbb{M}_i$ if $|\mathbb{M}_i| = 1$. Markets with a single prospective placement do not contribute to the matching likelihood.
- As $\psi \to 0$, $\mathcal{L}_M^{s_\upsilon,\psi}$ tends to the indicator function over the choice set, given the simulated errors. Formally,

$$\lim_{\psi \to 0} \mathcal{L}_M^{s_\upsilon,\psi}(M_i \,|\, \mathbf{\Omega}_i, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) = 1_{\mathcal{A}(M_i|\mathbf{\Omega}_i,\mathbf{Z}_i,\boldsymbol{\theta}_T,\boldsymbol{\theta}_M)}(\boldsymbol{\upsilon}^{s_\upsilon}),$$

where $\boldsymbol{\upsilon}^{s_\upsilon} = (\upsilon_M^{s_\upsilon})_{M \in \mathbb{M}_i}$ with $\upsilon_M^{s_\upsilon} = \sum_{c,h} M(c,h)\left[\varepsilon_{cy_h}^{s_\upsilon} + \eta_{x_c h}^{s_\upsilon}\right]$.

2. Integrate over $\mathbf{\Omega}_i$.

- To simulate $\boldsymbol{\omega}_{ch}$, let $\mathbf{\Gamma}_\omega$ be the Cholesky factor of $\mathbf{\Sigma}_\omega$. Draw fixed simulated values

$$\tilde{\boldsymbol{\omega}}_{ch}^{s_\omega} \sim \text{iid } N(0, I_{|\mathcal{R}_0|}), \quad s_\omega = 1, \dots, S_\omega,$$

for every $(c,h) \in C_i \times H_i$. Set $\boldsymbol{\omega}_{ch}^{s_\omega} = \mathbf{\Gamma}_\omega \tilde{\boldsymbol{\omega}}_{ch}^{s_\omega}$, and $\mathbf{\Omega}_i^{s_\omega} = \left(\boldsymbol{\omega}_{ch}^{s_\omega}\right)_{(c,h)\in C_i \times H_i}$.

- The conditional outcome likelihood $\mathcal{L}_{\mathbf{T},\mathbf{R}}(\mathbf{T}_i, \mathbf{R}_i \,|\, M_i, \mathbf{\Omega}_i^{s_\omega}, \mathbf{Z}_i, \boldsymbol{\theta}_T)$ has a closed-form.
- Define the simulated counterpart of the market-level likelihood $\mathcal{L}(M_i, \mathbf{T}_i, \mathbf{R}_i \,|\, \mathbf{Z}_i, \boldsymbol{\theta})$ as

$$\text{(SL)} \qquad \mathcal{L}^{S_\omega,S_\upsilon,\psi}(M_i, \mathbf{T}_i, \mathbf{R}_i \,|\, \mathbf{Z}_i, \boldsymbol{\theta}) = \frac{1}{S_\omega S_\upsilon} \sum_{s_\omega=1}^{S_\omega} \sum_{s_\upsilon=1}^{S_\upsilon} \mathcal{L}_M^{s_\upsilon,\psi}(M_i \,|\, \mathbf{\Omega}_i^{s_\omega}, \mathbf{Z}_i, \boldsymbol{\theta}_T, \boldsymbol{\theta}_M) \cdots$$
$$\cdots \mathcal{L}_{\mathbf{T},\mathbf{R}}(\mathbf{T}_i, \mathbf{R}_i \,|\, M_i, \mathbf{\Omega}_i^{s_\omega}, \mathbf{Z}_i, \boldsymbol{\theta}_T)$$

3. Add over markets and take logs:

- Finally, define:

$$\ell_n^{S_\omega,S_\upsilon,\psi}(\boldsymbol{\theta} \,|\, \mathbf{Z}) = \sum_{i=1}^{n} \log \mathcal{L}^{S_\omega,S_\upsilon,\psi}(M_i, \mathbf{T}_i, \mathbf{R}_i \,|\, \mathbf{Z}_i, \boldsymbol{\theta})$$

B Appendix: Matching Covariance

Claim 1. *The covariance matrix of the composite error* $\boldsymbol{\upsilon} = (\upsilon_M)_{M \in \mathbb{M}(C,H)}$ *is given by:*

$$cov(\upsilon_M, \upsilon_{M'}) = \sum_{c \in C} \sigma_\varepsilon(y_{M(c)}, y_{M'(c)}) + \sum_{h \in H} \sigma_\eta(x_{M(h)}, x_{M'(h)}). \tag{31}$$

Proof of Claim 1: Define $A(c,h) \equiv \varepsilon_{cy_h} + \eta_{x_c h}$. Note that

$$\begin{aligned} A(c,h)A(c',h') &= [\varepsilon_{cy_h} + \eta_{x_c h}][\varepsilon_{c' y_{h'}} + \eta_{x_{c'} h'}] \\ &= \varepsilon_{cy_h}\varepsilon_{c' y_{h'}} + \varepsilon_{cy_h}\eta_{x_{c'} h'} + \eta_{x_c h}\varepsilon_{c' y_{h'}} + \eta_{x_c h}\eta_{x_{c'} h'} \end{aligned}$$

From Assumption 3, it follows

$$\begin{aligned} \mathbb{E}A(c,h)A(c',h') &= 1\{c = c'\}\mathbb{E}\varepsilon_{cy_h}\varepsilon_{c' y_{h'}} + 1\{h = h'\}\mathbb{E}\eta_{x_c h}\eta_{x_{c'} h'} \\ &= 1\{c = c'\}\sigma_\varepsilon(y_h, y_{h'}) + 1\{h = h'\}\sigma_\eta(x_c, x_{c'}). \end{aligned}$$

Since $\mathbb{E}\upsilon_M = \mathbb{E}\upsilon_{M'} = 0$,

$$\begin{aligned} \text{cov}(\upsilon_M, \epsilon_{M'}) &= \mathbb{E}\upsilon_M \upsilon_{M'} \\ &= \mathbb{E}\left[\sum_{c,h} M(c,h)A(c,h)\right]\left[\sum_{c',h'} M'(c',h')A(c',h')\right] \\ &= \sum_{c,h}\sum_{c',h'} M(c,h)M'(c',h')\mathbb{E}A(c,h)A(c',h') \\ &= \sum_{c}\sum_{h,h'} M(c,h)M'(c,h')\sigma_\varepsilon(y_h, y_{h'}) \\ &\quad + \sum_{h}\sum_{c,c'} M(c,h)M'(c',h)\sigma_\eta(x_c, x_{c'}). \end{aligned}$$

Note that $\sum_{h,h'} M(c,h)M'(c,h')\sigma_\varepsilon(y_h, y_{h'}) = \sigma_\varepsilon(y_h, y_{h'})$ for h and h' such that $M(c,h) = M'(c,h') = 1$, which is equivalent to $\sigma_\varepsilon(y_{M(c)}, y_{M'(c)})$. Using a symmetric argument in the second term yields the desired expression:

$$\text{cov}(\upsilon_M, \epsilon_{M'}) = \sum_{c \in C} \sigma_\varepsilon(y_{M(c)}, y_{M'(c)}) + \sum_{h \in H} \sigma_\eta(x_{M(h)}, x_{M'(h)}).$$

Q.E.D.

Proof of Proposition 1: For an arbitrary market with choice set $\mathbb{M}(C,H)$, let

$$(32) \qquad \tilde{\upsilon}_M = \upsilon_M - \upsilon_{M_0} \quad \forall\, M \in \mathbb{M}(C,H) \setminus \{M_0\},$$

for some fixed $M_0 \in \mathbb{M}(C,H)$. Standard results (e.g. Train 2009) show that the covariance matrix of $\tilde{\boldsymbol{\upsilon}} \equiv (\tilde{\upsilon}_M)_{M\in\mathbb{M}(C,H)\setminus\{M_0\}}$ is identified up to a scale normalization. From (22), one can write the elements in the covariance matrix of $\tilde{\boldsymbol{\upsilon}}$ as follows:

$$(33) \qquad \begin{aligned} \text{cov}(\tilde{\upsilon}_{M'}, \tilde{\upsilon}_{M''}) &= \text{cov}(\upsilon_{M'} - \upsilon_{M_0}, \upsilon_{M''} - \upsilon_{M_0}) \\ &= \text{cov}(\upsilon_{M'}, \upsilon_{M''}) + \text{var}(\upsilon_{M_0}) - \text{cov}(\upsilon_{M'}, \upsilon_{M_0}) - \text{cov}(\upsilon_{M''}, \upsilon_{M_0}) \\ &= \sum_c \sigma_\varepsilon(y_{M'(c)}, y_{M''(c)}) + \sum_h \sigma_\eta(x_{M'(h)}, x_{M''(h)}) \\ &+ \sum_c \sigma_\varepsilon(y_{M_0(c)}) + \sum_h \sigma_\eta(x_{M_0(h)}) \\ &- \left[\sum_c \sigma_\varepsilon(y_{M_0(c)}, y_{M'(c)}) + \sum_h \sigma_\eta(x_{M_0(h)}, x_{M'(h)})\right] \\ &- \left[\sum_c \sigma_\varepsilon(y_{M_0(c)}, y_{M''(c)}) + \sum_h \sigma_\eta(x_{0M(h)}, x_{M''(h)})\right], \end{aligned}$$

$$(34) \qquad \begin{aligned} \text{var}(\tilde{\upsilon}_{M'}) &= \sum_c \sigma_\varepsilon(y_{M'(c)}) + \sum_h \sigma_\eta(x_{M'(h)}) \\ &+ \sum_c \sigma_\varepsilon(y_{M_0(c)}) + \sum_h \sigma_\eta(x_{M_0(h)}) \\ &- 2\left[\sum_c \sigma_\varepsilon(y_{M_0(c)}, y_{M'(c)}) + \sum_h \sigma_\eta(x_{M_0(h)}, x_{M'(h)})\right], \end{aligned}$$

where I write $\sigma_\eta(x) \equiv \sigma_\eta(x,x)$ and $\sigma_\varepsilon(y) \equiv \sigma_\varepsilon(y,y)$ to simplify notation.

First, I show how to identify the elements of the covariance matrix $\boldsymbol{\Sigma}_\eta$. Consider a market with three children, whose types are given by $x, x', x'' \in X$, and a single home, whose type is $y \in Y$. The set of feasible matchings in this market contains three matchings: $M_0 = (x,y)$, $M_1 = (x',y)$, and $M_2 = (x'',y)$, where I abuse notation and define the matching over the types of the children and homes. Using (33) and (34), one may see that the identified elements in the covariance matrix of $\tilde{\boldsymbol{\upsilon}}$ in this market are given by:

$$\sigma_1^* \equiv \frac{\text{cov}(\tilde{v}_{M_1}, \tilde{v}_{M_2})}{\text{var}(\tilde{v}_{M_1})} = \frac{\sigma_\eta(x', x'') + \sigma_\epsilon(y) + \sigma_\eta(x) - \sigma_\eta(x, x') - \sigma_\eta(x, x')}{2\sigma_\varepsilon(y) + \sigma_\eta(x) + \sigma_\eta(x') - 2\sigma_\eta(x, x')} \tag{35}$$

$$\sigma_2^* \equiv \frac{\text{var}(\tilde{v}_{M_2})}{\text{var}(\tilde{v}_{M_1})} = \frac{2\sigma_\varepsilon(y) + \sigma_\eta(x) + \sigma_\eta(x'') - 2\sigma_\eta(x, x'')}{2\sigma_\varepsilon(y) + \sigma_\eta(x) + \sigma_\eta(x') - \sigma_\eta(x, x')}. \tag{36}$$

Let $x = x_0$ and $y = y_0$, so

$$\sigma_1^* = \frac{\sigma_\eta(x', x'') + 1}{2 + \sigma_\eta(x')}. \tag{37}$$

Since $\sigma_\eta(x', x'') = \sigma_\eta(x')$ for $x'' = x'$, (37) identifies $\sigma_\eta(x')$ for an arbitrary $x' \in X$. Note that this implies that (37) also identifies $\sigma_\eta(x', x'')$ for arbitrary $x', x'' \in X$. This argument obtains since the distribution of η_h is independent of the specific market we consider, all of which are independent.

Second, I show how to identify the covariance matrix $\mathbf{\Sigma}_\varepsilon$. Consider a market with three children, whose types are given by $x, x', x'' \in X$, and two homes, whose types are $y, y' \in Y$. In this market, $\mathbb{M}(C, H)$ contains six matchings. Let M_0 be the matching that assigns placements (x, y) and (x', y'); M_1 the one that assigns (x, y) and (x'', y'), and M_2 the one assigning (x'', y) and (x', y'). Using (33), compute the following covariance:

$$\begin{aligned}
\text{cov}(\tilde{v}_{M_1}, \tilde{v}_{M_2}) &= \sigma_\varepsilon(y, y') + \sigma_\eta(x, x'') + \sigma_\eta(x', x'') \\
&\quad + \sigma_\varepsilon(y) + \sigma_\varepsilon(y') + \sigma_\eta(x) + \sigma_\eta(x') \\
&\quad - \left[\sigma_e(y) + \sigma_\eta(x) + \sigma_\eta(x', x'')\right] \\
&\quad - \left[\sigma_\varepsilon(y') + \sigma_\eta(x, x'') + \sigma_\eta(x')\right] \\
&= \sigma_\varepsilon(y, y').
\end{aligned}$$

Let M_3 be the matching assigning the placements (x', y) and (x, y'), and note that:

$$\text{cov}(\tilde{v}_{M_1}, \tilde{v}_{M_3}) = \sigma_\varepsilon(y') + \sigma_\varepsilon(y, y') + \sigma_\eta(x') + \sigma_\eta(x', x'') - 2\sigma_\eta(x, x''). \tag{38}$$

Hence, two elements of the covariance matrix of $\tilde{\upsilon}$ in this market, are given by:

$$\sigma_3^* = \frac{\text{cov}(\tilde{\upsilon}_{M_1}, \tilde{\upsilon}_{M_2})}{\text{var}(\tilde{\upsilon}_{M_1})} = \frac{\sigma_\varepsilon(y, y')}{2\sigma_\varepsilon(y') + \sigma_\eta(x') + \sigma_\eta(x'') - 2\sigma_\eta(x', x'')} \tag{39}$$

$$\sigma_4^* = \frac{\text{cov}(\tilde{\upsilon}_{M_1}, \tilde{\upsilon}_{M_3})}{\text{var}(\tilde{\upsilon}_{M_1})} = \frac{\sigma_\varepsilon(y') + \sigma_\varepsilon(y, y') + \sigma_\eta(x') + \sigma_\eta(x', x'') - 2\sigma_\eta(x, x'')}{2\sigma_\varepsilon(y') + \sigma_\eta(x') + \sigma_\eta(x'') - 2\sigma_\eta(x', x'')}. \tag{40}$$

Since $\mathbf{\Sigma}_\eta$ is identified, the previous two equations define the 2-by-2 system of equations:

$$\sigma_3^* = \frac{\sigma_\varepsilon(y, y')}{2\sigma_\varepsilon(y') + H} \tag{41}$$

$$\sigma_4^* = \frac{\sigma_\varepsilon(y') + \sigma_\varepsilon(y, y') + K}{2\sigma_\varepsilon(y') + H}, \tag{42}$$

where H and K are known constants. Identification of $\mathbf{\Sigma}_\varepsilon$ follows from noting that the above system of equations has a unique solution for $\sigma_\varepsilon(y')$ and $\sigma_\varepsilon(y, y')$, in terms of identified quantities. Q.E.D.