Designing the Menu of Licenses for Foster Care*

Altınok, Ahmet † MacDonald, Diana‡

April 2026

Abstract

This paper explores U.S. foster parent licensing, essential for placing foster children. We develop a theoretical matching model to study the optimal menu of licenses designed to match foster parents to children, under the presence of search frictions. Our findings highlight the following: (i) optimal allocation calls for a segregation of the market, (ii) simple transfer schedules achieve the purpose, (iii) complementarities do not ensure Positive Assortative Matching (PAM) in equilibrium. These are shown to be robust with respect to adverse selection, with suitable adaptations, to continuous ability types, as well as variations in search friction. Our results also suggest that the structure of the current licensing menu aligns with the optimal structure, but may fall short in screening.

***JEL* Classification:** C78, D47, D82

Keywords: Foster Care · Matching · Sorting · Search

*This work benefited from helpful suggestions and discussions with Marek Weretka, and participants at seminars at the ACM Conference on Equity and Access in Algorithms, Mechanisms and Optimization, ASSET Annual Meeting, ASU PhD Reunion Conference, Banco de México, CIDE, Conference on Mechanism and Institution Design, EPCS Annual Meeting, ITAM, LACEA Annual Meeting, Nazarbayev University, SING Annual Meeting, SSCW Annual Meeting, Stony Brook International Conference on Game Theory, Universidad de Chile, and Universidad de Guadalajara. The views and conclusions presented in this paper are exclusively of the authors and do not necessarily reflect those of Banco de México. All errors are our own.
Diana E. MacDonald gratefully acknowledges financial support from FONDECYT Postdoctoral, Chile (Grant No. 3200510), and the Institute for Research in Market Imperfections and Public Policy (MIPP, ICS13_002 ANID).

†Department of Economics, Nazarbayev University, ahmet.altinok@nu.edu.kz
‡Research Department, Banco de México, diana.macdonald@banxico.org.mx

1 Introduction

Foster care can be viewed as a two-sided matching market with heterogeneous children and parents, where foster parents have preferences over children, and child welfare agencies have preferences—on behalf of children—over foster parents.[1] These agencies facilitate the matching process through a menu of licenses offered to foster parents. Each license specifies the types of children a parent may foster and the corresponding transfers. In practice, children are grouped according to their level of care needs, and transfers vary across groups.

For example, foster parents in Arizona can choose between two licenses: *traditional* and *therapeutic*. In the former, foster parents can only foster children with standard needs, whereas in the latter they can foster not only children with standard needs, but also children with special needs. Transfers reflect this difference in care needs: foster parents receive around US$21 per day for a child with standard needs and about US$37 for a child with special needs. Similar *nested* licensing schemes appear in other foster care systems (e.g., Texas, Virginia), in child care (e.g., Vermont's Specialized Child Care endorsement), and in elder care (e.g., Arizona's tiered assisted living licenses), underscoring the broader applicability of our framework.

The menu of licenses shapes both the pool of potential matches and the financial incentives, as it determines which children each parent is eligible to care for. A poorly designed licensing menu risks creating mismatches that leave individuals in need vulnerable and unnecessarily inflate program costs. This paper develops a theoretical matching model to study the optimal menu of licenses designed for the U.S. foster care system.[2] We construct a two-sided matching model with heterogeneous agents (children differ in the level of care needed and parents differ in their ability to provide care), and a designer who coordinates match formation through a menu of contracts. The main innovation of our paper lies in introducing an *endogenous search friction* that varies with market size, an element not entirely within the designer's control. The analysis focuses on licenses that specify an allocation of parents across child submarkets and the corresponding transfers, as well as the sorting patterns that arise in equilibrium.

Our results suggest that the menu of licenses used in practice exhibits some of the properties of the optimal solution. First, we find that it is never optimal to randomly match all types of parents to all types of children, that is, the optimal allocation calls for segregation of the market. Second, we show that a simple transfer schedule achieves the purpose, that is, parents holding different licenses and providing care for the same

[1] See Appendix A for a detailed description of foster care in the U.S.

[2] Although motivated by foster care, the framework applies more broadly to other regulated matching environments — such as childcare, eldercare, and specialized health services — where licensing determines who meets whom and under what terms.

type of child can receive the same monetary transfer. Lastly, we find that complementarity between children's and parents' attributes is not sufficient to ensure that positive assortative matching (PAM) will arise in equilibrium. Thus, we provide sufficient conditions for the equilibrium sorting to exhibit PAM: either a *stronger* complementarity determined by the distribution of children' attributes, or a lower bound on the share of children with special needs.[3]

The model is as follows. There are two sides of the market populated by a continuum of agents: children and parents. Children are heterogeneous in the level of care needed—low-needs (x_1) or high-needs (x_2)—and parents are heterogeneous in their ability to provide care—low-ability (y_1) or high-ability (y_2). We start our analysis with this binary type space for parents, which we later extend to a continuum of types. We analyze the market under complete information; our results carry over, with suitable adaptations, when a parent's attribute is private information.We assume that children receive greater payoffs when matched than unmatched, and parents incur a cost when a match forms. The designer maximizes expected utility from children minus transfers to parents. We assume that the surplus of each match is nonnegative, thus profitable.[4] As in practice, we construct submarkets for each child's attribute, that is, there is a submarket populated by low-needs children and another submarket populated by high-needs children.

First, the designer announces and commits to a menu of licenses. A license specifies: **(1)** a randomization rule that determines the probability with which a parent is allocated into each submarket, and **(2)** a corresponding transfer when a match forms. After observing the menu, each parent chooses a license. Next, the randomization device is realized and parents are allocated across submarkets determining endogenously the parents-to-children ratio (market tightness) for each submarket.[5] Lastly, within each submarket, meetings take place, matches are formed, and transfers occur. We introduce a search friction by assuming that meetings are not certain, that is, the probability of a child (parent) meeting a parent (child) is represented by a meeting technology which is a function of the market tightness. Thus, when parents choose a license that guarantees allocation to a specific submarket with probability one, they are certain about the type of child they will be matched with but remain uncertain about whether the match will occur. However, if a parent holds a license that allocates them to either submarket with a strictly positive probability, they face uncertainty not only about the type of child they will be matched with but also whether a match will take

[3]Ideally, we would empirically test our theoretical predictions on optimal sorting patterns, but this is not feasible due to limited data on parents' attributes.

[4]In our framework, surplus of a match is a cost-net benefit function whose arguments are parent's ability and child's level of care needed.

[5]We use the language of allocation across submarkets, but it can also be interpreted as the weight the designer assigns to a specific parent holding a particular license to provide care for one type of child or another.

place.

Our search friction assumption is motivated by the fact that child welfare agencies do not act as matchmakers but instead define feasible matches through a menu of licenses and guidelines. In practice, social workers are responsible for contacting parents about a specific child in a decentralized manner. Thus, the randomization mechanism can be interpreted as a guideline: for example, if a parent is deemed 'better' suited to care for low-needs children than high-needs children, the system would aim to allocate that parent to the first submarket with a higher probability. Furthermore, market tightness captures the level of congestion in the market, while the meeting technology accounts for the frictions arising from the decentralized nature of the matching process which strongly depends on the congestion. Therefore, our model incorporates endogenous market tightness via a meeting technology, creating trade-offs between match quality and matching probability.

It is important to emphasize that the search friction is a crucial element of our model, as it introduces non-trivial effects on the analysis.[6] Specifically, when a mass of type-y parents is reallocated from one submarket to another, three key effects occur: **(i)** *Surplus effect:* This represents the change in total expected surplus of the market. **(ii)** *Congestion effect:* The change in market tightness in the submarket where parents are reallocated, leading to a thicker market. **(iii)** *Decongestion effect:* The change in market tightness in the submarket from which parents are reallocated, resulting in a thinner market. These effects not only add complexity to the analysis but also enrich the predictions of our model.

We begin by examining the case with complete information and derive results for both super- and sub-modular surplus functions. In this section, we focus on the case of super-modularity, while the discussion of sub-modularity is deferred to the main body of the paper.

First, we find that it is optimal to use at least one type of parents exclusively in one submarket, that is, to segregate parents as much as possible.[7] The way the segregation result is established is as follows: We perturb candidate allocations of the two ability types along "iso-tightness" lines. Such perturbations change the ability composition in each submarket without changing the overall matching probabilities. Since the objective function of the planner–the expected surplus–is linear in these perturbations, the planner should move along an iso-tightness line to an edge of the space of feasible allo-

[6] Our paper is closely related to the work of Damiano and Li (2007), which explores how a monopoly matchmaker sorts agents into exclusive meeting places for random pairwise matching. A key assumption in their framework is that all agents have a constant match probability of one, abstracting from size effects that could influence matching probabilities based on market scale. In contrast, our paper incorporates market size effects through the search friction assumption, illustrating how market size influences matching probabilities.

[7] Formally, if the optimal randomization rule is interior for type-y parents, then it is a corner solution for type-y' parents, where y and y' are distinct.

cations. This result speaks to the optimality of the nested hierarchy property exhibited in the licenses used in practice. That is, the traditional license allocates parents exclusively into the submarket of children with standard needs, whereas the therapeutic license can potentially allocate parents into both submarkets of children with standard and special needs; in the case of Arizona.

Second, we show that **simple transfers** are always part of an optimal menu; in particular, the transfer need not vary with licenses—a uniform *per child price* suffices. More specifically, parents with different licenses caring for the same type of child can receive the same transfer at the optimum. This arises under the nested structure of the equilibrium allocation, which segments the market: the corner license has its transfer pinned down by its participation constraint, while the license that spans both submarkets has a single binding constraint that leaves one degree of freedom. This flexibility allows the transfer in the shared submarket to be aligned across licenses. This observation helps rationalize the pricing rule used by Arizona's social welfare agency—US$21 per day for children with standard needs and US$37 for children with special needs, regardless of the license held. However, the *magnitude* of transfers must still depend on the child's attributes and on market primitives such as the distribution of agents.

Third, we show that supermodularity of the surplus is neither necessary nor sufficient on its own for the optimal sorting to exhibit PAM. However, when the surplus is supermodular and sufficiently strong relative to the frictions in the market, this combination is sufficient to generate PAM. In our framework, the randomization device determines who can match with whom, and we use it to define sorting patterns: sorting exhibits PAM (NAM) if y_2-parents are allocated to submarket x_2 with a greater (smaller) probability than y_1-parents.[8] For a frictionless environment with a supermodular surplus function, it is well known that positive assortative matching maximizes total welfare (Becker, 1973). When search frictions are introduced, however, the expected surplus—computed using the meeting technologies in each submarket—need not be supermodular even if the underlying surplus function is. Finally, we show that PAM can be ensured at the optimum by imposing a lower bound on the fraction of type-x_2 children together with supermodularity. Intuitively, type-y_2 parents are more valuable in any submarket, so the designer prefers to allocate them to the more profitable and *thicker* submarket x_2.

In this context, one might imagine that the child welfare agency could screen foster parents using observable characteristics such as race, marital status, education, employment status, or income. Under such a scenario, our complete-information anal-

[8]One can equivalently define sorting through a matching correspondence, as is standard in the literature, and say that sorting exhibits PAM if the matching correspondence is a lattice as in Shimer and Smith (2000). Since the randomization device carries more information than the correspondence, our notion is more general: any feasible unequal allocation of parents in our setting exhibits either PAM or NAM, but not both, unlike Shimer and Smith (2000).

ysis would suffice. However, the literature shows that these observable traits do not predict the likelihood of fostering higher-needs children, whereas the type of license foster parents hold does.[9] This motivates our first extension, which relaxes the assumption of full observability of parental attributes, and demonstrates that our main results still hold. Next, while retaining the high- and low-need classification of children, we generalize the parental attribute space to a continuum to better capture heterogeneity. Finally, we perform comparative statics on the meeting technology, assessing robustness to variations in the decentralized search process across U.S. states. Taken together, these extensions confirm that segregation, the optimality of simple transfer schedules, and the insufficiency of complementarities for PAM all carry over.

Literature review. The market-design literature in foster care is relatively narrow. Existing work has largely focused on improving information flows, search processes, or platform design rather than on how institutional features—such as licensing menus —structure feasible matches. A first strand studies information and matching platforms. Slaugh et al. (2016) analyze the Pennsylvania Adoption Exchange and show that centralized information and match-recommendation tools can improve the matching process and facilitate adoption placements. Robinson-Cortés (2021) documents how cross-regional placements and thicker matching environments affect placement stability, highlighting the role of geographic integration in child-welfare outcomes. Dierks et al. (2024) develop a dynamic search-and-matching model that emphasizes the importance of caseworker-directed search, and MacDonald (2024) provides empirical and theoretical evidence on transition patterns between foster and adoption placements. A second strand focuses on mechanism design and allocation rules. Dierks et al. (2024) propose incentive-compatible platforms tailored to children with disabilities, Baron et al. (2024) show that redesigned case-assignment mechanisms can improve child-welfare outcomes without harming investigators, and Highsmith (2024) develops dynamic placement mechanisms that outperform traditional sequential procedures.

Our paper contributes by analyzing a foundational but understudied institutional feature: the *menu of licenses*. Licenses determine eligibility, sort parents into distinct submarkets, and shape market tightness —yet the existing literature typically treats the set of feasible matches as exogenous. We show instead that licensing menus endogenously generate submarkets, influence parents' entry and search decisions, and interact with search frictions to determine equilibrium sorting patterns. We further establish that the allocation maximizing child welfare —requiring segregation across license types— can be implemented through simple, incentive-compatible transfer schedules.

[9] Cox et al. (2011) finds no significant association between foster mothers' observable characteristics and the likelihood of fostering children with emotional and behavioral problems.

In doing so, our analysis connects to the broader literature on submarket formation and endogenous market segmentation in directed-search environments (Peters, 1997; Montgomery, 1991; Chade and Smith, 2006; Chade et al., 2017), and to the literature on matching with contracts and transfers (Kelso and Crawford, 1982; Hatfield and Milgrom, 2005).

Our framework also departs from the classical directed–search literature, in which submarkets arise endogenously from agents' posting decisions or strategic signaling (Montgomery, 1991; Peters, 1997; Moen, 1997). In these models, the planner cannot directly shape market segmentation; submarkets are an equilibrium outcome rather than a design choice. In foster care, by contrast, the designer explicitly *creates* the submarket structure through the licensing menu, which determines eligibility and the distribution of placement probabilities across child–need types. Because submarkets are designer-created rather than agent-created, licensing menus become a policy instrument that shapes incentives, market tightness, and the sorting patterns that emerge in equilibrium.

This institutional perspective also connects our analysis to the broader search-and-matching literature that studies how frictions and complementarities interact with submarket formation. Menzio and Shi (2010a,b) formalize the role of directed search, submarkets, and market tightness; Shi (2001) shows that supermodularity alone is insufficient for Positive Assortative Matching (PAM) in directed-search environments; and Eeckhout and Kircher (2010) characterize the complementarity thresholds required for assortative matching. Similarly, Shimer and Smith (2000) and Smith (2006) show that under two-sided random search, PAM requires strong conditions such as log-supermodularity. Consistent with these insights, we demonstrate that in foster care, search frictions make strong complementarities necessary—but not sufficient—for PAM to arise, and that the designer's ability to engineer submarkets through licensing menus fundamentally shapes the resulting matching patterns.

Empirical relevance. Although the analysis is theoretical, the model yields clear and testable predictions that can be examined using administrative foster-care microdata —particularly in states where detailed licensing and child-needs information exceed what is publicly available in AFCARS.[10] First, the model predicts that changes in licensing rules should alter *sorting and segregation patterns* across submarkets, measurable through segregation indices as in Card et al. (2013). Second, it predicts deviations from PAM when complementarities are weak or search frictions are high, producing observable mismatches between parental attributes and child needs. Third, the model implies that *payment differentials across license types* should be larger in markets with

[10] AFCARS is a federally mandated data collection system. All fifty US states and the District of Columbia are required to collect data on all children in foster care and all children adopted from foster care

greater heterogeneity in child needs or parent characteristics, linking reimbursement schedules to underlying type dispersion. Fourth, *market tightness*—the availability of licensed foster homes relative to children needing placement—should adjust to licensing reforms, generating shifts in the number of active homes across license types and in time-to-placement. These empirical patterns can be studied using segregation tools from Card et al. (2013) and assignment-based evaluation frameworks from the matching and allocation literature (e.g., Abdulkadiroğlu et al. (2005, 2011, 2017)), as well as insights from the broader literature on how licensing and regulatory regimes shape market structure Autor et al. (2006, 2010).

Organization of the paper. The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents the analysis under complete information. In Section 4, we develop three extensions: (i) we relax the assumption of full observability of parental attributes, (ii) we generalize the parental attribute space to a continuum to capture richer heterogeneity, and (iii) we perform comparative statics on the meeting technology to assess robustness to differences in decentralized search processes across U.S. Lastly, Section 5 concludes. Appendix A presents an overview of foster care in the U.S., and all omitted proofs are in Appendices B, C, and D. The online supplement contains proofs and examples omitted from the main text.

2 Model

One side of the market is populated by a continuum of children who differ in an observable attribute $x \in X = \{x_1, x_2\}$, where x_1 denotes a low-needs child (without a disability), x_2 denotes a high-needs child (with a disability), and $x_2 > x_1$. The fraction of children with low needs is $f(x_1) \in [0, 1]$, whereas the fraction with high needs is $f(x_2) = 1 - f(x_1)$. We refer to the set of children with attribute x as submarket x. The other side of the market is populated by a continuum of parents who are heterogeneous in their ability to provide care. In particular, y_1 denotes parents with low-ability, y_2 denotes parents with high-ability, and $y_2 > y_1$.[11] The fraction of parents with low-ability is $g(y_1) \in [0, 1]$, and that with high-ability is $g(y_2) = 1 - g(y_1)$.

Matches are formed between children and parents on a one-to-one basis.[12] There is a **designer** who facilitates the matching process by offering a menu of licenses to parents. A license $\mathcal{L}$ is represented by a pair (λ, τ), where $\lambda : X \to [0, 1]$ is a randomization device that determines the probability with which a parent is allocated to submarket x, and $\tau : X \to \mathbb{R}$ represents a transfer between the designer and the parent if a match

[11] Section 4 expands the type space of parents to a continuum.

[12] According to Gibbs and Wildfire (2007), the average occupancy rate is 1.5 children per home, indicating that assuming one-to-one matches aligns with the empirical evidence.

with a child of type x is formed..[13] Throughout the paper, we restrict attention to the menu of licenses with the following features: (i) allocations are non-wasteful, that is, $\sum_{x \in X} \lambda(x) = 1$, and (ii) parents have limited liability, that is, $\tau(x) \geq 0$ for all $x \in X$.

Figure 1 represents two examples of randomization devices under separate licenses. Parents holding license $\mathcal{L}$ are allocated to submarket x_1 with probability one, and to submarket x_2 with probability zero. Analogously, parents holding license $\mathcal{L}'$ are in submarkets x_1 and x_2 with probabilities $1/4$ and $3/4$, respectively.

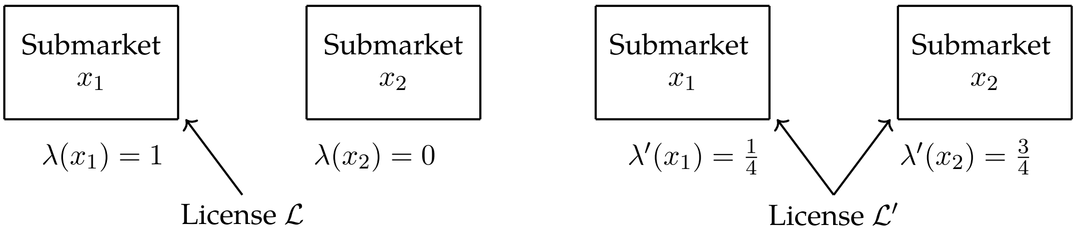

Figure 1: Examples of Randomization Devices

All agents are risk-neutral. The designer maximizes children's welfare net of transfers. Payoffs for unmatched agents are normalized to zero. When a child x and a parent y form a match, the child receives payoffs according to a real-valued function $u(x, y)$, and the parent incurs a cost of providing care according to a real-valued function $c(x, y)$.[14]

Assumption 1. *(a) for all (x, y), $u(x, y) \geq 0$, $c(x, y) \geq 0$ and $u(x, y) - c(x, y) \geq 0$, (b) $u(x, y)$ is increasing in y, and (c) $c(x, y)$ is increasing in x and decreasing in y.*

Assumption 1(a) reflects the following: children are better-off placed with a foster parent than being unmatched, parents incur a cost when providing care for a child, and all matches are profitable. Assumption 1(b) states that children prefer high- to low-ability parents. Finally, Assumption 1(c) implies that parents incur a smaller cost when matched to low-needs children than to high-needs children, and high-ability parents incur a smaller cost when providing care than low-ability parents. We present the timeline in what follows:

1. First, the designer announces and commits to a menu of licenses:
$\{\mathcal{L}^k\}_{k=1}^2 \equiv \left\{ \{ (\lambda^k(x_i), \tau^k(x_i)) \}_{i=1}^2 \right\}_{k=1}^2$.

2. In the complete-information benchmark, the designer assigns each parent a license from the menu; in the private-information extension, each parent chooses a license from the menu. Let $\sigma^y \in \{\mathcal{L}^1, \mathcal{L}^2\}$ denote the license held by a type-y parent. Then, the

[13] Alternatively, $\lambda(x)$ can be interpreted as the probability with which a parent is considered to provide care for a type-x child.

[14] The parents' cost function can be interpreted as a net cost function, which captures the balance between the benefits of providing care for a child and the associated costs.

allocation of parents $\left\{\{\lambda^k(x_i)\}_{i=1}^2\right\}_{k=1}^2$ across submarkets is realized.[15]

3. Next, children and parents in each submarket meet stochastically. The meeting technology can be described in terms of the parents-to-children ratio (*market tightness*). The market tightness of each submarket $x \in X$, denoted by θ_x, is equal to:

$$\theta_x = \frac{\sum_{k=1}^2 \lambda^k(x) \sum_{j=1}^2 h_k(y_j)}{f(x)}.$$

where $h^k(y)$ denotes the endogenous mass of parents $y \in \{y_1, y_2\}$ choosing license k. A child x meets a parent according to a *technology* $\pi^c(\theta_x)$ where $\pi^c : \mathbb{R}_+ \to [0,1]$ is a strictly increasing and strictly concave function such that $\pi^c(0) = 0$.[16,17]

4. Finally, when a child x and a parent y meet, a match (x, y) is formed and transfers take place according to $\left\{\{\tau^k(x_i)\}_{i=1}^2\right\}_{k=1}^2$.

Designer's Problem: The designer aims to maximize children's welfare while minimizing the transfers. We start by specifying the objective function of the designer. Let $\mathcal{L} \equiv \left\{\left\{\left(\lambda^k(x_i), \tau^k(x_i)\right)\right\}_{i=1}^2\right\}_{k=1}^2$ be an arbitrary menu of licenses. A child x receives utility $u(x, y_j)$ when she matches with a parent y_j. Notice that parent y_j might hold either contract, thus the net utility when a child x matches with parent y_j under contract k is $u(x, y_j) - \tau^k(x)$. Now, conditional on a meeting taking place, the probability that child x has met a parent y_j holding license k is equal to: $\lambda^k(x)h^k(y_j) / \sum_{k=1}^{2} \left(\lambda^k(x) \sum_{j=1}^{2} h^k(y_j)\right)$. Thus, the net expected utility in each submarket x, conditional on a meeting taking place, is:

$$W(x) = \frac{\sum_{k=1}^{2} \left[\sum_{j=1}^{2} \left[u(x, y_j) - \tau^k(x)\right] \cdot \lambda^k(x) \cdot h^k(y_j)\right]}{\sum_{k=1}^{2} \lambda^k(x) \cdot \left[\sum_{j=1}^{2} h^k(y_j)\right]}.$$

Then, the designer's problem is:

$$\max_{\left\{\left\{\left(\lambda^k(x_i), \tau^k(x_i)\right)\right\}_{i=1}^2\right\}_{k=1}^2} \left\{ \sum_{i=1}^{2} \pi^c(\theta_{x_i})\, W(x_i)\, f(x_i) \right\} \quad \text{subject to:} \tag{1}$$

$$[\text{FC}] \quad \tau^k(x) \geq 0 \text{ and } \lambda^k(x) \geq 0 \text{ for all } (k, x), \text{ and } \sum_{i=1}^{2} \lambda^k(x_i) = 1 \text{ for all } k = 1, 2.$$

$$[\text{MT}] \quad \theta_x = \frac{1}{f(x)} \cdot \sum_{k=1}^{2} \left[\lambda^k(x) \sum_{j=1}^{2} h^k(y_j)\right], \text{ for all } x.$$

$$[\text{PC}] \quad \sum_{i=1}^{2} \left[\tau^k(x_i) - c(x_i, y_k)\right] \lambda^k(x_i) \pi^p(\theta_{x_i}) \geq 0, \text{ for all } k = 1, 2.$$

[15] Under a license with interior randomization, a parent faces ex-ante uncertainty over the child's type, but any realized match involves only one child type ex post.

[16] Similarly, a parent meets a child x with probability $\pi^p(\theta_x)$ where $\pi^p : \mathbb{R}_+ \to [0,1]$ is a strictly decreasing and convex function such that $\pi^p(\theta_x) = \pi^c(\theta_x)/\theta_x$ and $\pi^p(0) = 1$. This relationship ensures that the probability of a child meeting a parent is consistent with the probability of a parent meeting a child, by equating the expected number of parents' meetings to children's.

[17] The search friction assumption highlights the decentralized matching process in the U.S. foster care system, where market congestion plays a key role. As market tightness increases, parents are less likely to find children, while children are more likely to find parents. These frictions also account for the possibility that congestion may prevent some matches from forming.

where [FC] are the feasibility constraints specifying restrictions over each $\lambda^k(x)$ and $\tau^k(x)$. The restriction [MT] corresponds to the market tightness (parents-to-children ratio) in each submarket. The constraints [PC] are the participation constraints guaranteeing that each parent y_j receives a weakly higher expected payoff from holding license $k = j$ than from remaining unmatched.

Sorting Patterns: Now, we define a matching correspondence and establish sorting patterns using the randomization device specified via the menu: $\{\lambda^1(x_i), \lambda^2(x_i)\}_{i=1}^2$.

Definition 1. *A* ***matching correspondence*** *is a map $\mu : Y \to X$ such that $x \in \mu(y_{\mathbf{k}})$ if and only if $\lambda^{\mathbf{k}}(x) > 0$. Moreover, if $\lambda^2(x_2) \geq \lambda^1(x_2)$ then the sorting exhibits* ***Positive Assortative Matching (PAM)****. Analogously, if $\lambda^1(x_2) \geq \lambda^2(x_2)$ then the sorting exhibits* ***Negative Assortative Matching (NAM)****.*

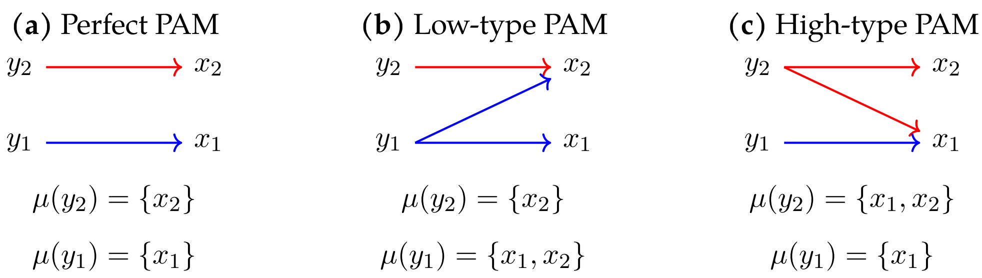

Figure 2: Examples of Positive Assortative Matching (PAM)

We are interested not only in establishing properties that ensure monotone sorting, but also in characterizing the optimal menu of licenses offered by the designer. As a result, our notion of monotone sorting is as follows: We say that sorting exhibits high-type PAM (NAM) if type-y_2 parents are allocated to both submarkets, while type-y_1 parents are allocated only to submarket x_1 (x_2). Analogously, low- type PAM (NAM) follows.

Figure 2 presents examples illustrating our concept of monotone sorting patterns. In Panel 2b, for instance, y_2-parents are allocated into submarket x_2 with probability one and y_1-parents are allocated into both submarkets with strictly positive probability, thus $1 = \lambda^2(x_2) \geq \lambda^1(x_2) \in (0, 1)$. Note that the randomization device in Panel 2c can represent the menu of licenses used in practice, as outlined in the introduction.[18]

3 Equilibrium Analysis: Complete Information

In this section, we examine the optimal menu of licenses and analyze sorting patterns that arise in equilibrium under complete information. We focus on symmetric equilibria in which parents of the same type are assigned the same license. First, note that by

[18]In this case, parents holding license 1 are restricted to fostering only low-needs children i.e., $\mu(y_1) = \{x_1\}$, while parents with license 2 are eligible to foster both types of children i.e., $\mu(y_2) = \{x_1, x_2\}$, reflecting the nested structure highlighted previously.

incorporating the [PC] constraints into the objective function in Equation (1) reduces the designer's problem to:

$$\max_{\{\lambda^k(x_1),\lambda^k(x_2)\}_{k=1}^2} \left\{ \sum_{i=1}^{2} \pi^p\big(\theta_{x_i}\big) \left[\sum_{k=1}^{2} \Big(\underbrace{u(x_i, y_k) - c(x_i, y_k)}_{S(x,y)} \Big) \lambda^k(x_i)\; g(y_k) \right] \right\} \tag{2}$$

subject to [FC] and [MT]. Let θ_1 and θ_2 denote θ_{x_1} and θ_{x_2}, respectively. In addition, let $S(x, y) \equiv u(x, y) - c(x, y)$ denote the **surplus of a match** (x, y) which is increasing in y by Assumption 1.

Lemma 1 (**Segregation**)**.** *For at least one of the licenses, the optimal randomization rule (allocation) yields a corner solution.*

Please see Appendix B.1 for the proof. Lemma 1 states that it is never optimal to allocate parents of both ability types across the two submarkets. Whether positive assortative or negative assortative, the optimal matching is always on an "edge" of the space of feasible allocations.

Now, to characterize the optimal randomization rule, we follow a nonstandard technique due to the presence of corner solutions. We start with an arbitrary interior allocation and examine whether the designer can increase total expected welfare by simply reallocating parents across submarkets. Formally, for each (x, k), let $\lambda^k(x)$ be an arbitrary-feasible interior probability that generates a total welfare equal to:

$$\begin{aligned} W\big(\lambda^1(x_1), \lambda^2(x_1)\big) &= \pi^p(\theta_1) \cdot \Big[g(y_1)\; \lambda^1(x_1)\; S(x_1, y_1) + \big(1 - g(y_1)\big)\; \lambda^2(x_1)\; S(x_1, y_2) \Big] \\ &\quad + \pi^p(\theta_2) \cdot \Big[g(y_1)\; \big(1 - \lambda^1(x_1)\big)\; S(x_2, y_1) + \big(1 - g(y_1)\big)\; \big(1 - \lambda^2(x_1)\big)\; S(x_2, y_2) \Big] \end{aligned}$$

After perturbing $\lambda^1(x_1)$ by ε_1 and $\lambda^2(x_1)$ by ε_2 where

$$\varepsilon_2 \equiv -\frac{\varepsilon_1 g(y_1)}{1 - g(y_1)},$$

so that ensuring that market tightness remains constant, the change in welfare is equal to:

$$\begin{aligned} \Delta_W &= W\big(\lambda^1(x_1) + \varepsilon_1, \lambda^2(x_1) + \varepsilon_2\big) - W\big(\lambda^1(x_1), \lambda^2(x_1)\big) \\ &= \varepsilon_1\; g(y_1)\; \underbrace{\Big(\pi^p(\theta_2)\; \big[S(x_2, y_2) - S(x_2, y_1)\big] - \pi^p(\theta_1)\; \big[S(x_1, y_2) - S(x_1, y_1)\big] \Big)}_{Z^{CI}(\theta_1)} \end{aligned}$$

Since $\theta_2 = \frac{1 - f(x_1)\theta_1}{1 - f(x_1)}$, $Z^{CI}(\cdot)$ can be viewed as a function of θ_1. Notice, the designer can always increase total welfare by changing $\big(\lambda^1(x_1), \lambda^2(x_1)\big)$ such that the market tightness remains constant. The optimal allocation of parents can be characterized by $Z^{CI}(\theta_1)$, which represents the expected difference in gains between children x_2 and x_1 from being matched to a high-ability parent as opposed to a low-ability parent. Moreover, the sign of $Z^{CI}(\theta_1)$ determines the equilibrium sorting. Let $\bar{\theta}_1$ be such that $Z^{CI}(\bar{\theta}_1) = 0$, then the following result holds:

Proposition 1. *Let θ_1^* be the equilibrium market tightness.* ***(i)*** *If $\theta_1^* > \bar{\theta}_1$ then the equilibrium sorting exhibits PAM.* ***(ii)*** *If $\theta_1^* < \bar{\theta}_1$ then the equilibrium sorting exhibits NAM.* ***(iii)*** *$\theta_1^* = \bar{\theta}_1$ is never optimal.*

If the equilibrium market tightness θ_1^* is such that $Z^{CI}(\theta_1^*)$ is positive then PAM arises in equilibrium. To see this, note that $Z^{CI}(\theta_1)$ is increasing in θ_1, that is, the change in welfare increases as θ_1 increases. Thus, $Z^{CI}(\theta_1) > 0$ for any $\theta_1 > \bar{\theta}_1$. Therefore, when $Z^{CI}(\theta_1)$ is positive, we can pick $\varepsilon_1 > 0$, increasing the share of y_1-parents allocated in submarket x_1 and, vice versa for y_2-parents, effectively moving towards the edge of the feasible allocations. Intuitively, a high θ_1^* translates into a small probability of a parent meeting a child in submarket x_1. Since y_2-parents generate a greater surplus, it would be optimal to minimize the probability with which they remain unmatched. Thus, the designer allocates y_2-parents in submarket x_2, leading to PAM. Analogously, NAM follows. See Appendix B.2 for the proof.

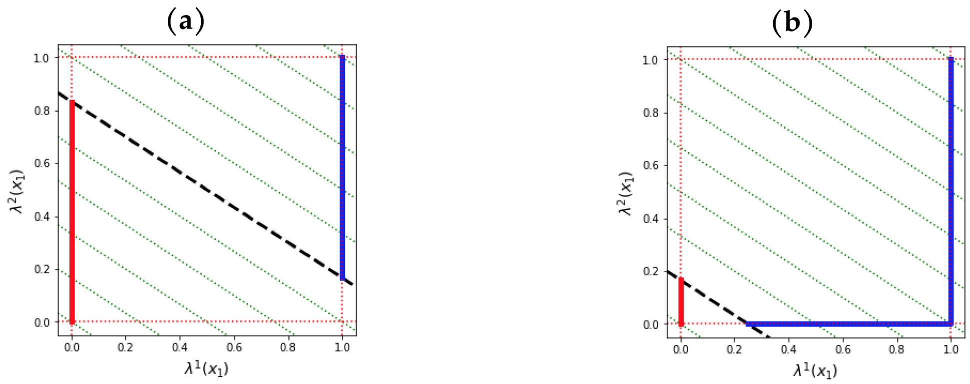

Figure 3: Illustration of PAM and NAM given $Z^{CI}(\bar{\theta}_1)$

Figure 3 illustrates Lemma 1 and Proposition 1. In each box, the x- and y-axis correspond to the probability with which parents holding license 1 and 2 are allocated into submarket x_1, respectively. Thus, every point inside the box $\big(\lambda^1(x_1), \lambda^2(x_1)\big)$ is a feasible allocation of parents. Yet, note that by Lemma 1 only the points at the borders can be an equilibrium. In addition, each black-dashed line corresponds to the values of $\big(\lambda^1(x_1), \lambda^2(x_1)\big)$ such that $Z^{CI}(\bar{\theta}_1) = 0$, each blue line shows the feasible allocations that can be an equilibrium when $Z^{CI}(\theta_1) > 0$ (above the black-dashed line), and each red line shows the feasible allocations that can be an equilibrium when $Z^{CI}(\theta_1) < 0$ (below the black-dashed line). In Panel 3a, the equilibrium candidates are along the vertical blue line and vertical red line. In the former, allocations are such that $\lambda^2(x_2) \geq \lambda^1(x_2) = 0$, which corresponds to high-type PAM. In the latter, allocations are such that $1 = \lambda^1(x_2) \geq \lambda^2(x_2)$, which corresponds to high-type NAM. Analogously, in Panel 3b, the equilibrium candidates are along the red and the blue lines.

Now, we are interested in establishing sufficient conditions for PAM and NAM to arise in equilibrium. The optimal allocation segregates the market and implements PAM only when surplus is supermodular, and strong enough relative to the search friction. Corollary 1 follows directly from Proposition 1.[19]

Corollary 1 (**Sufficient conditions for PAM and NAM**). ***(i)*** *If* $\frac{S(x_2,y_2)-S(x_2,y_1)}{S(x_1,y_2)-S(x_1,y_1)} \geq \frac{1}{\pi^p(1/f(x_2))}$ *holds, then the equilibrium sorting exhibits PAM.* ***(ii)*** *If* $\frac{S(x_1,y_2)-S(x_1,y_1)}{S(x_2,y_2)-S(x_2,y_1)} \geq \frac{1}{\pi^p(1/f(x_1))}$ *holds, then the equilibrium sorting exhibits NAM.*

For Corollary 1(i), notice that $Z^{CI}(\theta_1)$ reaches its minimum value at $\theta_1 = 0$, implying that $\pi^p(0) = 1$ and $\theta_2 = \frac{1}{f(x_2)}$. Thus, we ensure PAM by imposing that the minimum value of $Z^{CI}(\theta_1)$ is positive. Observe that (i) requires a super-modular surplus function since the right-hand side is greater than 1. Moreover, the greater the left-hand side of (i) is, the stronger the supermodularity is. Thus, *strong* supermodularity on the surplus function dominates the adverse effect of the search friction, and becomes sufficient to induce PAM at the optimum.[20]

Alternatively, one can think of the inequality (i) as a lower bound over the share of children with high-needs to ensure PAM in equilibrium. Intuitively, by imposing a lower bound on the share of type-x_2 children we ensure that the market is thick enough for the more desirable type-y_2 parents, that is, the probability of meeting a child in submarket x_2 is bounded below. This is in line with the literature in dynamic search and matching, which imposes stronger complementarity conditions to ensure that more desirable partners have incentives to wait for a more desirable partner from the other side of the market.[21] Similar arguments and intuition follows for (ii). The proof is provided in Appendix B.3.

Next, we study the optimal transfer schemes. By fixing the optimal allocations

[19] Corollary 1 ensures that PAM or NAM will arise in equilibrium, but it does not specify whether we will observe either low-type PAM (NAM), high-type PAM (NAM), or perfect PAM (NAM). Refer to Appendix B.1 in the online supplement for a detailed characterization.

[20] By *strong* super-modularity on the surplus function we mean: $[S(x_2,y_2) - S(x_2,y_1)] \cdot \pi^p\left(\frac{1}{f(x_2)}\right) \geq S(x_1,y_2) - S(x_1,y_1)$ which introduces a constraint that is sensitive to the underlying distribution and the specific meeting technology. Similarly, the condition of *strong* sub-modularity on the surplus function follows.

[21] Shimer and Smith (2000) and Smith (2006) analyze a dynamic two-sided matching setting with heterogeneous agents, random search and complete information. The former paper assumes that utility is fully transferable and establishes as a sufficient condition not only supermodularity on the value of a match $f(x,y)$ where x and y are the agent's attributes, but also supermodulariy on $log\ f_x$ and $log\ f_{xy}$. The latter paper assumes that utility is strictly non-transferable and establishes as sufficient conditions monotonicity and log-supermodularity in $f(x,y)$. In both papers, these conditions ensure that, in the search process, high-partners do not settle for a low-partner but instead wait for the arrival of a high-partner. This is in the same spirit as our condition: we are also making sure that the payoffs received from matching high-types together compensate for the adversary effect of search frictions.

$\{\lambda^{k*}(x_1), \lambda^{k*}(x_2)\}_{k=1}^2$ from Equation 2, the designer solves the following:

$$\min_{\{\tau^k(x_1),\tau^k(x_2)\}_{k=1}^2} \left\{ \sum_{i=1}^{2} \pi^p(\theta_i^*) \sum_{k=1}^{2} \tau^k(x_i)\lambda^{k*}(x_i)\, g(y_k) \right\}$$

subject to [FC], [MT], and [PC] from Equation 1. The following proposition states that the transfer scheme is characterized by [PC]:

Proposition 2 (**Simple Transfers**). *At any equilibrium $\{\lambda^{k*}(x_1), \lambda^{k*}(x_2)\}_{k=1}^2$, any feasible transfer schedule for which the participation constraints hold with equality is an equilibrium.*

Recall that the optimal allocation of at least one type of parent is a corner solution, in which case the transfer can be trivially pinned down. For example, suppose that the equilibrium sorting exhibits perfect PAM, that is, y_1-parents are allocated into submarket x_1 with probability one, while y_2-parents are allocated into submarket x_2 with probability one. Then, the optimal transfer scheme is $\tau^{1*}(x_1) = c(x_1, y_1)$ and $\tau^{2*}(x_2) = c(x_2, y_2)$. That is, parents receive exactly the cost of providing care, as is the current practice. See Appendix B.4 for the proof.

In the case of an interior allocation, the optimal transfer scheme is not unique. As an example, suppose that the equilibrium sorting exhibits high-type PAM, that is, y_1-parents are allocated into submarket x_1 with probability one, while y_2-parents are allocated into both submarkets with strictly positive probability. Note, this is similar to the example of Arizona discussed in the introduction where low-needs children can be fostered by parents holding any of the two licenses, and high-needs children can only be fostered by parents holding one particular license. Here, the optimal transfer scheme is $\tau^{1*}(x_1) = c(x_1, y_1)$, $\tau^{2*}(x_1) \geq 0$, and $\tau^{2*}(x_2) = c(x_2, y_2) - [\tau^{2*}(x_1) - c(x_1, y_2)]\frac{\pi^p(\theta_1^*)\lambda^{2*}(x_1)}{\pi^p(\theta_2^*)\lambda^{2*}(x_2)}$. Now, as in practice, suppose that parents who provide care in the same market receive the same transfer, that is, $\tau^{1*}(x_1) = \tau^{2*}(x_1) = c(x_1, y_1)$. In this case, the optimal transfer for parent y_2 in submarket x_2 would be the following:

$$\tau^{2*}(x_2) = c(x_2, y_2) - \left[c(x_1, y_1) - c(x_1, y_2)\right]\frac{\pi^p(\theta_1^*)\lambda^{2*}(x_1)}{\pi^p(\theta_2^*)\lambda^{2*}(x_2)}$$

In what follows, we briefly discuss what the theoretical findings imply over foster care in the U.S..

3.1 Implications for foster care in the U.S.

There are several implications for the design of licensing systems in foster care.

First, the agency should design the menu to separate parents as much as possible. Specifically, parental abilities should be allocated to distinct submarkets—exclusively where feasible. When overlap is unavoidable, it is preferable to assign a single parental ability across multiple submarkets rather than to combine multiple abilities within several submarkets. In this regard, the nested hierarchy adopted in some existing state-level menus provides a useful benchmark consistent with this principle.

Second, the agency should assign parents with higher abilities—taking into account overall surplus generation—to larger submarkets. Doing so minimizes the costs associated with search frictions. For example, when the submarket for children with special needs is large, it becomes more likely that the optimal menu exhibits Positive Assortative Matching (PAM) rather than Negative Assortative Matching (NAM), and the converse holds when the submarket is small.

Finally, a **simple transfer schedule** is sufficient. At the optimum, parents with different licenses caring for the same type of child can receive the same transfer. This result stems directly from the nested structure of the equilibrium allocation, which implies that markets are segmented. In contrast, with a fully interior allocation, such a simple transfer schedule would not satisfy both [PC] conditions simultaneously. Given that the designer has the freedom to choose any randomization mechanism, it is not at all obvious that simple transfers can always be part of an equilibrium. Additionally, equilibrium transfers depend on further market characteristics, including the number of children, the number of parents, the allocation itself, and the meeting technology.

4 Extensions

Here, we extend our complete information model in three key directions and present the corresponding findings. The segregation result, the insufficiency of supermodularity for PAM and the optimality of simple transfer schedules are robust to adverse selection and extend, with suitable adaptations, to settings with continuous ability types. We also analyze the sensitivity of our results by conducting comparative statics on the meeting technology.

4.1 Private Information

In this section, we analyze the case where a parent's ability is private information by incorporating incentive compatibility constraints into the problem specified in Equation (1):

$$[\text{IC}] \quad \sum_{i=1}^{2} \left[\tau^k(x_i) - c(x_i, y_k)\right] \lambda^k(x_i)\pi^p(\theta_{x_i}) \geq \sum_{i=1}^{2} \left[\tau^{k'}(x_i) - c(x_i, y_k)\right] \lambda^{k'}(x_i)\pi^p(\theta_{x_i}),\ \forall k, k' = 1, 2$$

Recall that high-ability parents incur a smaller cost when providing care for any child than low-ability parents do. Thus, high-ability parents receive a smaller expected transfer under the optimal menu specified for the complete information setting, regardless of the sorting pattern or equilibrium allocation. Hence, high-ability parents have incentives to mimic low-ability parents in the presence of private information. Therefore, the [PC] for low-ability parents and the [IC] for high-ability parents must be binding in equilibrium. Plugging them into the objective function in Equation (1), the designer's problem reduces to:

$$\max_{\{\lambda^k(x_1),\lambda^k(x_2)\}_{k=1}^2} \left\{ \sum_{i=1}^{2} \pi^p(\theta_i) \left[\sum_{k=1}^{2} \Big(\underbrace{u(x_i,y_k) - c(x_i,y_k)}_{S(x,y)} \Big) \lambda^k(x_i)\, g(y_k) \right] \right.$$
$$\left. - \Big[c(x_1,y_1) - c(x_1,y_2)\Big] \lambda^1(x_1)\pi^p(\theta_1) g(y_2) - \Big[c(x_2,y_1) - c(x_2,y_2)\Big] \lambda^1(x_2)\pi^p(\theta_2) g(y_2) \right\} \tag{3}$$

subject to [FC], [MT], and [IC] for low-ability parents.[22,23]

As one can see, in the objective function, extra terms appear in the second line due to information frictions. Looking closely, these terms correspond to the expected gain that a high-ability parent obtains by mimicking low-ability parents.

Remark 1. *The incentive compatibility conditions are satisfied if and only if:*

$$\Big(\lambda^2(x_2) - \lambda^1(x_2)\Big) \underbrace{\Big(\pi^p(\theta_2)\big[c(x_2,y_1) - c(x_2,y_2)\big] - \pi^p(\theta_1)\big[c(x_1,y_1) - c(x_1,y_2)\big]\Big)}_{C(\theta_1)} \geq 0$$

where $C(\cdot)$ represents the expected cost difference between caring for child x_2 and child x_1 when comparing a high-ability parent to a low-ability parent. Note that, $C(\theta_1)$ increases with θ_1. Furthermore, $\lambda^2(x_2) - \lambda^1(x_2)$ is non-negative if the given allocation exhibits PAM.

With this remark in hand, we start by showing that Lemma 1 and Proposition 1 presented in the previous section carry over to the private information (see Appendix C.1 and C.2). To do so, we initially examine the relaxed version of the designer's problem (Equation 3) by dropping [IC] for low-ability parents. After that we introduce a condition, using Remark 1, for both to be satisfied. Following the perturbation argument over iso-tightness line, as we did previously, the term that characterizes the optimal allocation of parents across submarkets becomes:

[22]Notice, [PC] for low-ability parents and [IC] for high-ability parents imply [PC] for high-ability parents, see proof of Proposition 4 in Appendix C.4.

[23]For the complete information case, the assumption $\sum_{x\in X} \lambda(x) = 1$ does not play a role in our results: if we relax it to $\sum_{x\in X} \lambda(x) \leq 1$, at the optimum this inequality will still be binding. In the private information setting, the optimum could change if we relax this equality: the designer might find it optimal to leave some foster parents out of the market to mitigate the incentives of mimicking. However, we believe that our assumption is reasonable considering that foster care exhibits a shortage of foster parents, who have to pass a rigorous assessment to be accepted to participate in the market. Thus, imposing that the system would like to employ all available parents is in line with the child welfare agencies objectives. In addition, relaxing this assumption would make the problem intractable for the private information case.

$$
\begin{aligned}
Z^{PI}(\theta_1) = \pi^p(\theta_2)\Big[S(x_2,y_2) - S(x_2,y_1) + \frac{g(y_2)}{g(y_1)}\Big(c(x_2,y_1) - c(x_2,y_2)\Big)\Big] \\
- \pi^p(\theta_1)\Big[S(x_1,y_2) - S(x_1,y_1) + \frac{g(y_2)}{g(y_1)}\Big(c(x_1,y_1) - c(x_1,y_2)\Big)\Big] \quad (4)
\end{aligned}
$$

Note that $Z^{PI}(\theta_1)$ is analogous to $Z^{CI}(\theta_1)$, adjusted by the cost due to the information friction. Recall,

$$
\begin{aligned}
Z^{CI}(\theta_1) = \pi^p(\theta_2)\Big(\big[u(x_2,y_2) - c(x_2,y_2)\big] - \big[u(x_2,y_1) - c(x_2,y_1\big]\Big) \\
- \pi^p(\theta_1)\Big(\big[u(x_1,y_2) - c(x_1,y_2)\big] - \big[u(x_1,y_1) - c(x_1,y_1)\big]\Big) \quad (5)
\end{aligned}
$$

A couple of significant insights are worth highlighting. First, if there are no high-ability parents (i.e., $g(y_1) = 1$), then there is no screening problem and Equations (4) and (5) coincide. Second, under private information, the information friction amplifies the *cost difference* term in the virtual surplus. In particular, in (4) the expression

$$
c(x_i, y_1) - c(x_i, y_2)
$$

is weighted by $g(y_2)/g(y_1) = 1/g(y_1) - 1$, which increases its magnitude whenever $g(y_1) < 1$. This is the sense in which information frictions magnify the cost component—they scale the *difference* in costs, rather than the level $c(x, y)$ itself. Third, the size of this amplification decreases as the share of low-ability parents increases.

As in Section 3, the sign of $Z^{PI}(\cdot)$ at the equilibrium θ_1 determines the equilibrium sorting (see Proposition C.1). Now, we present sufficient conditions for monotone sorting under private information, analogous to Corollary 1:[24]

Corollary 2 (**Sufficient Conditions for PAM and NAM**). *The equilibrium sorting exhibits*

(*i*) *PAM if* $\frac{S(x_2,y_2)-S(x_2,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_2,y_1)-c(x_2,y_2)]}{S(x_1,y_2)-S(x_1,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_1,y_1)-c(x_1,y_2)]} \geq \frac{1}{\pi^p(1/f(x_2))}$ *and* $\frac{c(x_2,y_1)-c(x_2,y_2)}{c(x_1,y_1)-c(x_1,y_2)} \geq \frac{1}{\pi^p(1/f(x_2))}$.

(*ii*) *NAM if* $\frac{S(x_1,y_2)-S(x_1,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_1,y_1)-c(x_1,y_2)]}{S(x_2,y_2)-S(x_2,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_2,y_1)-c(x_2,y_2)]} \geq \frac{1}{\pi^p(1/f(x_1))}$ *and* $\frac{c(x_1,y_1)-c(x_1,y_2)}{c(x_2,y_1)-c(x_2,y_2)} \geq \frac{1}{\pi^p(1/f(x_1))}$.

Unlike the complete information case, the surplus function is not sufficient to elicit the equilibrium sorting pattern under the presence of private information. Here, we need to take into account the cost of a match as well as the distribution of parents. The first condition of Corollary 2(i) ensures that the allocation that maximizes the objective function exhibits PAM, and the second condition guarantees that the allocation is implementable—incentive compatible.

Note, the second condition requires $c(x, y)$ to be a *strong* sub-modular function.[25] That is, the difference between the cost of providing care for a child x_2 and a child x_1

[24]See Appendix B.2 in the online supplement for a detailed characterization.

[25]By *strong* sub-modularity on the cost function we mean: $[c(x_2, y_1) - c(x_2, y_2)] \cdot \pi^p\left(\frac{1}{f(x_2)}\right) \geq c(x_1, y_1) - c(x_1, y_2)$ which introduces a constraint that is sensitive to the underlying distribution and the specific meeting technology. Similarly, the condition of *strong* super-modularity on the cost function follows.

must be greater for low-ability parents than for high-ability parents. A sub-modular cost function implies that the informational rents paid to high-ability parents are lower under PAM than NAM. In other words, it is cheaper for the designer to motivate high-ability parents to report truthfully while inducing PAM. Analogous for 2(ii).

Next, motivated by the fact that the child welfare agency may not have the distribution of parents' attributes, we establish conditions that do not depend on this primitive:

Corollary 3 (**Sufficient Conditions for PAM and NAM**). *The equilibrium sorting exhibits*

(*i*) *PAM if* $\frac{S(x_2,y_2)-S(x_2,y_1)}{S(x_1,y_2)-S(x_1,y_1)} \geq \frac{1}{\pi^p(1/f(x_2))}$ *and* $\frac{c(x_2,y_1)-c(x_2,y_2)}{c(x_1,y_1)-c(x_1,y_2)} \geq \frac{1}{\pi^p(1/f(x_2))}$.

(*ii*) *NAM if* $\frac{S(x_1,y_2)-S(x_1,y_1)}{S(x_2,y_2)-S(x_2,y_1)} \geq \frac{1}{\pi^p(1/f(x_1))}$ *and* $\frac{c(x_1,y_1)-c(x_1,y_2)}{c(x_2,y_1)-c(x_2,y_2)} \geq \frac{1}{\pi^p(1/f(x_1))}$.

Corollary 3 follows directly from Corollary 2. In item (i), we require $S(x,y)$ to be a strong super-modular function as in the complete information case, and the condition of strong sub-modularity in $c(x,y)$ to ensure incentive compatibility.

Next, we examine the problem under more relaxed conditions, specifically when the conditions outlined in Corollaries 2 and 3 are not satisfied. Let $C(\theta_1') = 0$ and $Z^{PI}(\theta_1'') = 0$ for some θ_1' and θ_1''. Thus, $C(\theta_1) > 0$ if and only if $\theta_1 > \theta_1'$, and $Z^{PI}(\theta_1) > 0$ if and only if $\theta_1 > \theta_1''$.

Proposition 3. *Let θ_1^{**} denote the equilibrium market tightness derived from the solution $\{\lambda^{k^{**}}(x_1)\}_{k=1}^2$ in Equation 3.* **(*i*)** *If $\theta_1^{**} \geq \max\{\theta_1', \theta_1''\}$ then $\{\lambda^{k^{**}}(x_1)\}_{k=1}^2$ is a solution to Equation 1 which exhibits PAM.* **(*ii*)** *If $\theta_1^{**} \leq \min\{\theta_1', \theta_1''\}$ then $\{\lambda^{k^{**}}(x_1)\}_{k=1}^2$ is a solution to Equation 1 which exhibits NAM.* **(*iii*)** *Otherwise, $\{\lambda^{k^{**}}(x_1)\}_{k=1}^2$ and the corresponding induced θ_1^{**} do not solve Equation 1.*

Proposition 3 is analogous to Proposition 1. It essentially states that the solution to the relaxed problem forms an equilibrium for the more constrained problem if the resulting market tightness in submarket-x_1 is either sufficiently high or low, relative to specific thresholds determined by the meeting technology, surplus, and cost functions: $\max\{\theta_1', \theta_1''\}$ and $\min\{\theta_1', \theta_1''\}$.

Next, we analyze the equilibrium transfers under private information. By fixing the optimal allocations $\{\lambda^{k**}(x_1), \lambda^{k**}(x_2)\}_{k=1}^2$ from Equation 3, the designer solves the following: $\min_{\{\tau^k(x_1),\tau^k(x_2)\}_{k=1}^2} \left\{ \sum_{i=1}^2 \pi^p\left(\theta_i^{**}\right) \sum_{k=1}^2 \tau^k(x_i)\lambda^{k**}(x_i)\, g(y_k) \right\}$ subject to [FC], [PC] and [IC] from Equation 1. Here, the [PC] for low-ability parents, and the [IC] for high-ability parents determine the equilibrium transfer scheme. Formally:

Proposition 4 (**Simple Transfers**). *Fix an equilibrium allocation $\{\lambda^{k**}(x_1), \lambda^{k**}(x_2)\}_{k=1}^2$. Any feasible transfer schedule that satisfies both of the PC for y_1-type parents and the IC for y_2-type parents with equality is an equilibrium provided that one of the following conditions holds:* **(*i*)** *$\lambda^{2**}(x_2) \geq \lambda^{1**}(x_2)$ and $c(x,y)$ is strong sub-modular, or* **(*ii*)** *$\lambda^{1**}(x_2) \geq \lambda^{2**}(x_2)$ and $c(x,y)$ is strong super-modular*

Note that conditions in Corollaries 2 and 3 are also sufficient for Proposition 4. In particular, Corollary 2(i) or Corollary 3(i) ensure two things: (1) the cost function $c(x, y)$ is strong sub-modular, and (2) the randomization device exhibits PAM, that is, $\lambda^{2**}(x_2) \geq \lambda^{1**}(x_2)$. Thus, conditions in Corollaries 2 or 3 satisfy the conditions in Proposition 4(i). An analogous argument applies to item (ii). See Appendix C.4 for the proof. Notice, simple transfer schedules are still part of an optimal menu, despite the information frictions.

We conclude this section with an example by incorporating private information.

Example 1. (***PAM fails despite a supermodular surplus function***) *Let the environment have the following specifications:* $f(x_1) = 0.8$, $g(y_1) \in (0, 1)$, $S(x_2, y_2) = 191$, $S(x_1, y_2) = 201$, $S(x_2, y_1) = 40$, $S(x_1, y_1) = 51$, *and that* $\pi^p(\theta) = {}^1/_{1+\theta}$. *Moreover, suppose that the cost function is super-modular with the following values:* $c(x_2, y_2) = 15$, $c(x_1, y_2) = 1$, $c(x_2, y_1) = 20$ *and* $c(x_1, y_1) = 15$.[26]

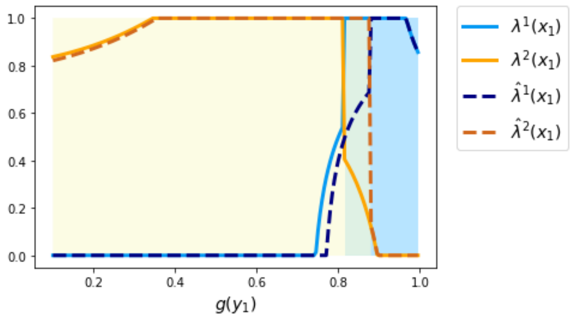

Figure 4: Randomization Device - Private & Complete Information

Figure 4 shows that the optimal randomization devices for both complete $(\lambda^k(x_1), 1-\lambda^k(x_1))$, *and private information* $(\hat{\lambda}^k(x_1), 1-\hat{\lambda}^k(x_1))$ *scenarios are remarkably similar. Note that* $\lambda^1(x_1)$ *and* $\lambda^2(x_1)$ *closely resemble* $\hat{\lambda}^1(x_1)$ *and* $\hat{\lambda}^2(x_1)$, *respectively.*[27] *However, when* $g(y_1)$ *lies approximately in* $(0.8, 0.9)$, *the equilibrium sorting pattern is PAM under complete information, whereas it is NAM with private information.*

To build intuition, note first that under these parameter values only NAM can satisfy the incentive compatibility constraints; a separating PAM allocation is not feasible. The numerical comparison below therefore serves solely to illustrate why information rents move in a way that disfavors PAM. Under complete information, consider the equilibrium menu that achieves perfect sorting, so that $\tau^k(x) = c(x, y_k)$. *Suppose this menu induces* ***NAM***. *A type-*y_2 *parent then has an incentive to misreport as type-*y_1 *in order to reach* x_2 *rather than* x_1, *yielding an (ex-post) deviation gain*

$$\tau^1(x_2) - c(x_2, y_2) = 5.$$

[26]Notice, the cost function here guarantees the existence of a separating menu of licenses under NAM, whereas any equilibrium exhibiting PAM does not screen parents.

[27]This similarity arises due to the values of the surplus and cost functions. If the values of the cost function were to increase, it would lead to a notable disparity in the optimal randomization rule between the complete and private information settings.

*If instead the same menu were used to implement **PAM**, the analogous misreporting incentive becomes*

$$\tau^1(x_1) - c(x_1, y_2) = 14,$$

which is strictly larger. Hence, type-y_2 parents would have stronger *incentives to mimic type-y_1 under PAM than under NAM. This illustrates why, for the present parameter values, NAM is the only sorting pattern that can be supported under private information, even though PAM is optimal under complete information for some values of $g(y_1)$. This intuition aligns with the counterpart to Corollary 2(i). For a detailed analysis, see Appendix A of the online supplement.* □

4.2 Continuous ability types

Motivated by the fact that a parent's ability to provide care might be a continuous variable, we now assume that parents differ in $y \in [\underline{y}, \bar{y}] \equiv Y \subset \mathbb{R}_+$, which follows a continuous and differentiable cumulative distribution function (CDF) $G(y)$ with a strictly positive probability density function (PDF) $g(y)$. We uphold our regularity conditions: $S(x, y)$ is increasing in y, while $c(x, y)$ is increasing in x and decreasing in y.

Now, a license for a type-y parent is $\mathcal{L}(y) \equiv \{(\lambda(x_i, y), \tau(x_i, y))\}_{i=1}^2$. Thus, conditional on a meeting taking place, the probability that a child of type x meets a parent of type y is $\lambda(x,y)g(y) / \int_{\underline{y}}^{\bar{y}} \lambda(x,y)g(y)\, dy$.

Hence, the net expected utility in each submarket x, conditional on a meeting, is $W(x) = \left(\int_{\underline{y}}^{\bar{y}} [u(x,y) - \tau(x,y)]\, \lambda(x,y)g(y)\, dy\right) / \left(\int_{\underline{y}}^{\bar{y}} \lambda(x,y)g(y)\, dy\right)$. The designer's problem under complete information becomes:

$$\max_{\left(\{(\lambda(x_i,y),\tau(x_i,y))\}_{i=1}^2\right)_{y\in Y}} \left\{\sum_{i=1}^{2} \pi^c(\theta_i)\, W(x_i)\, f(x_i)\right\} \tag{6}$$

subject to [FC] and [PC] as defined earlier, and [MT]: $\theta_i = 1/f(x_i) \cdot \int_{\underline{y}}^{\bar{y}} \lambda(x_i, y)g(y)\, dy$, for $i = 1, 2$.

After incorporating the [PC] equations into the objective function, and using the [MT] along with the relationship $\pi^c(\theta)/\theta = \pi^p(\theta)$, the designer's problem reduces to:

$$\max_{\left(\{\lambda(x_i,y)\}_{i=1}^2\right)_{y\in Y}} \left\{\sum_{i=1}^{2} \pi^p(\theta_i) \cdot \int_{\underline{y}}^{\bar{y}} S(x_i, y)\lambda(x_i, y)g(y)\, dy\right\} \tag{7}$$

It is easy to see that the randomization device $\lambda(x, y)$ is independent of whether interim or ex-post participation constraints are satisfied (see Corollary B.1). Furthermore, the segregation result from the two-type case (see Lemma 1) extends to the current environment, albeit with a caveat:

Lemma 2. *Any **interior** randomization $\lambda(x, y)$ is **sub-optimal**.*

Please see Appendix D.1 for the proof. Lemma 2 states that for each parent y, the randomization device will take a value of either zero or one. Notably, unlike in the complete-information case, a parent will foster only one type of child here. Therefore, the nested structure of licenses described in the introduction does not appear in this case.

The proof follows the two–type case. Begin with an interior randomization $\lambda(x_1, y) \in (0, 1)$ so that $\lambda(x_2, y) = 1 - \lambda(x_1, y)$ by the [FC]. Define a perturbation $\varepsilon : Y \to \mathbb{R}$ constructed *piecewise* over intervals on which $\hat{Z}^{CI}(y \mid \boldsymbol{\theta})$ is monotone, and impose $\int_I \varepsilon(y)g(y)\,dy = 0$ on each such interval I. Let $\tilde{\lambda}(y) = \big(\lambda(x_1, y) - \varepsilon(y),\ \lambda(x_2, y) + \varepsilon(y)\big)$, which preserves market tightness. The welfare change is

$$\Delta_W = \int_{\underline{y}}^{\overline{y}} \hat{Z}^{CI}(y \mid \boldsymbol{\theta})\,\varepsilon(y)g(y)\,dy, \text{where } \hat{Z}^{CI}(y \mid \boldsymbol{\theta}) := \pi^p(\theta_2)S(x_2, y) - \pi^p(\theta_1)S(x_1, y).$$

By choosing ε to increase where $\hat{Z}^{CI}$ increases and decrease where $\hat{Z}^{CI}$ decreases, each monotonicity interval contributes positively to Δ_W, so any interior randomization is sub-optimal. [28]

Definition 2. *If $\lambda(x_2, y)$ is non-decreasing (non-increasing) in y, then sorting exhibits PAM (NAM).*

Given an equilibrium market tightness (θ_1^*, θ_2^*), there exists a threshold parent attribute such that all parents with ability below this threshold are allocated to one submarket, while all parents with ability above it are allocated to the other submarket. This result reflects a complete segregation of the market.

Proposition 5. *Let (θ_1^*, θ_2^*) be an equilibrium market tightness.*

(*i*) *If $\hat{Z}^{CI}(y \mid \boldsymbol{\theta}^*)$ is increasing in y, then the equilibrium exhibits PAM:*

$$\lambda^*(x_2, y) = \begin{cases} 0, & y \le \hat{y}^{PAM}, \\ 1, & y > \hat{y}^{PAM}, \end{cases} \qquad \theta_1^* = \frac{G(\hat{y}^{PAM})}{1 - f(x_2)}, \quad \theta_2^* = \frac{1 - G(\hat{y}^{PAM})}{f(x_2)}.$$

(*ii*) *If $\hat{Z}^{CI}(y \mid \boldsymbol{\theta}^*)$ is decreasing in y, then the equilibrium exhibits NAM:*

$$\lambda^*(x_2, y) = \begin{cases} 1, & y \le \hat{y}^{NAM}, \\ 0, & y > \hat{y}^{NAM}, \end{cases} \qquad \theta_1^* = \frac{1 - G(\hat{y}^{NAM})}{1 - f(x_2)}, \quad \theta_2^* = \frac{G(\hat{y}^{NAM})}{f(x_2)}.$$

[28] Appendix D.1 provides the explicit piecewise construction. The analogy between $\partial_y \hat{Z}^{CI}$ and the discrete Z^{CI} is heuristic; the argument only relies on the sign and monotonicity of $\hat{Z}^{CI}$. We thank an anonymous referee for pointing to it.

See Appendix D.2 for the proof.[29] In what follows, we establish sufficient conditions for monotone sorting. To this end, let $S(x,y)$ be continuous and differentiable over Y, and let $S_y(x,\cdot)$ denote the partial derivative of $S(x,y)$ with respect to y.

Corollary 4.

(i) *If* $S_y(x_2,\grave{y})/S_y(x_1,\grave{y}) \geq \frac{1}{\pi^p(1/f(x_2))}$ *where* $\grave{y} := \arg\min_{y\in Y} \pi^p\left(1/1-f(x_1)\right)S_y(x_2,y) - S_y(x_1,y)$*, then the equilibrium sorting exhibits PAM. That is,* $\lambda^*(x_2,y) = 0$ *for all* $y \leq \hat{y}^{PAM}$ *and* $\lambda^*(x_2,y) = 1$ *for all* $y > \hat{y}^{PAM}$ *with:*

$$\hat{y}^{PAM} := \arg\max_{\hat{y}\in Y} \pi^p\left(G(\hat{y})/f(x_1)\right)\int_{\underline{y}}^{\hat{y}} S(x_1,y)g(y)dy + \pi^p\left(1-G(\hat{y})/1-f(x_1)\right)\int_{\hat{y}}^{\overline{y}} S(x_2,y)g(y)dy.$$

(ii) *If* $S_y(x_1,\acute{y})/S_y(x_2,\acute{y}) \geq \frac{1}{\pi^p(1/f(x_1))}$ *where* $\acute{y} := \arg\max_{y\in Y} S_y(x_2,y) - \pi^p\left(1/f(x_1)\right)S_y(x_1,y)$*, then the equilibrium sorting exhibits NAM. That is,* $\lambda^*(x_2,y) = 1$ *for all* $y \leq \hat{y}^{NAM}$ *and* $\lambda^*(x_2,y) = 0$ *for all* $y > \hat{y}^{NAM}$ *with:*

$$\hat{y}^{NAM} := \arg\max_{\hat{y}\in Y} \pi^p\left(1-G(\hat{y})/f(x_1)\right)\int_{\hat{y}}^{\overline{y}} S(x_1,y)g(y)dy + \pi^p\left(G(\hat{y})/1-f(x_1)\right)\int_{\underline{y}}^{\hat{y}} S(x_2,y)g(y)dy.$$

See Appendix D.3 for a proof. Corollary 4 (i) implies that $S(x,y)$ is supermodular. This suggests that our condition is slightly stronger than the standard supermodularity of the surplus function required to ensure PAM, as in the two-type case. Analogously, a sufficient condition—Corollary 4 (ii)—for NAM follows. Furthermore, these conditions result in a segregation of the market, with one key distinction from the two-type case: no parent is allocated across both submarkets. Specifically, a threshold type-y parent emerges, dividing the type space into two distinct partitions, each allocated to a separate submarket. This contrasts with environments involving a finite number of parent types. An immediate implication is that if the conditions for monotone sorting are satisfied, the designer's problem simplifies to selecting the threshold type-y parent that partitions the type space. This choice is effectively equivalent to determining the market tightness.

Remark 2. *The model and analysis can be readily extended to accommodate an arbitrary environment with* $X = \{x_1, x_2, \ldots, x_n\}$ *where* $n \geq 2$ *while* $y \sim Y$ *with a nonzero PDF* $g(\cdot)$*. In such a setting, whenever* $\lambda(x,y) \in (0,1)$ *for some* $x \in \{x_i, x_j\}$ *over some non-zero measure* $Y' \subseteq Y$*, a similar perturbation—between submarkets* i *and* j *over the type-space* Y' *without altering MT at each submarket—yields the following change in the welfare:*

$$\Delta_W = \int_{y\in Y'} \left[\pi^p(\theta_{x_i})S(x_i,y) - \pi^p(\theta_{x_j})S(x_j,y)\right]\varepsilon(y)g(y)dy.$$

[29]The proof relies on the monotonicity of $\hat{Z}^{CI}(y \mid \boldsymbol{\theta}^*)$ in y. Supermodularity (respectively, submodularity) of the underlying bivariate function $\pi^p(\boldsymbol{\theta}^*)S(x,y)$ is a *sufficient* condition for this monotonicity, but supermodularity is not a property of the univariate mapping $y \to \hat{Z}^{CI}(y \mid \boldsymbol{\theta}^*)$. We thank the referee for this clarification.

Thus, the steps outlined above can be easily followed to replicate the analysis and derive characterizations analogous to Proposition 5 as well as Corollary 4.

Lastly, the transfers can be easily determined by the [PC] since parents are allocated to exactly one submarket. Specifically, $\tau^*(x_i, y) = c(x_i, y)$ if $\lambda^*(x_i, y) = 1$ for all (x_i, y). Thus, in this case, we observe that paying parents exactly the cost of providing care, as mentioned in the Arizona example, constitutes an equilibrium. However, it is important to note that this outcome is optimal only when licenses are not nested.

4.3 The Effect of Search Friction

In this section, we analyze the role of meeting technology in modeling search frictions within the allocation process. Specifically, we investigate how changes in search technology—whether advancements or setbacks—affect our sorting results. This analysis is motivated by the observed differences in the effectiveness of child welfare agencies in matching children with suitable foster families.

Formally, we define what constitutes an improvement in search technology, concentrating, without loss of generality, on its application to parents. Recall that $\pi^p(\cdot)$ is a strictly decreasing and strictly convex function bounded by $\pi^p(0) = 1$ and $\lim_{\theta\to\infty} \pi^p(\theta) = 0$.[30] Let $\mathbf{\Pi}^p$ be the set of all such bounded, strictly decreasing, and strictly convex functions. For any $\pi^p, \hat{\pi}^p \in \mathbf{\Pi}^p$, we say that $\hat{\pi}^p$ is an **improved technology** compared to π^p if $\partial\hat{\pi}^p(\theta)/\partial\theta > \partial\pi^p(\theta)/\partial\theta$ for any finite θ. Note, this also implies $\hat{\pi}^p(\theta) > \pi^p(\theta)$. If $\hat{\pi}^p$ is an improved technology of π^p, we denote we denote $\hat{\pi}^p \rhd \pi^p$. Now, the following partially characterizes the equilibrium sorting as meeting technology improves:

Proposition 6. *Suppose $S(x, y)$ is supermodular (submodular). If the equilibrium sorting is PAM (NAM) with some meeting technology π^p, then it remains PAM (NAM) for any $\hat{\pi}^p$ such that we denote $\hat{\pi}^p \rhd \pi^p$.*

5 Concluding Remarks

This paper analyzes the foster care system in the U.S. as a two-sided matching market wherein one side consists of children who are heterogeneous in the level of care needed, and the other side consists of parents who differ in their ability to take care of a child. We solve for the optimal menu of licenses which specifies an allocation of parents across submarkets of children as well as the corresponding transfers, in the presence of search and information frictions.

[30]Moreover, recall that $\pi^c(\theta)/\theta \equiv \pi^p(\theta)$. And thus, for any other meeting technology for parents $\tilde{\pi}^p(\theta)$, it has to be the case that $\tilde{\pi}^p(\theta) \cdot \theta = \tilde{\pi}^c(\theta)$ is strictly increasing and concave with the following bounds $\tilde{\pi}^c(0) = 0$ and $\lim_{\theta\to\infty} \tilde{\pi}^c(\theta) = 1$.

The paper establishes three key results that hold regardless of the information frictions: **(i)** it is not optimal to mix multiple types of parents into multiple submarkets of children, **(ii)** a single price for a type of a child at each license is indeed part of an optimal menu, and **(iii)** super-modularity and sub-modularity of the surplus of a match are neither sufficient nor necessary conditions for the optimal sorting to exhibit PAM and NAM, respectively. The former two rationalize the nested nature of the menu of licenses offered by various states in the U.S.. The latter has implications for the optimal allocation of parents: even if the surplus shows complementarity (substitutability) in child and parent's attributes, allocating parents into submarkets such that the sorting exhibits PAM (NAM) is not necessarily optimal due to search frictions.

We also make inferences once information friction is introduced: as the share of low-type parents increases, the allocation of parents approaches the first-best (complete information). This is because high-type parents mimic the low-type parents to receive a greater expected transfer. As a result, the designer pays information rents to high-type parents to overcome such incentives. The smaller the share of high-type parents, the less costly information friction becomes. However, if the proportion of high-type parents is big enough, then not only the allocation diverges from the first-best, but also the optimal sorting may reverse.

Lastly, we analyze the sensitivity of our results by introducing a continuous attribute space for parents and briefly discuss the implications of expanding the discrete attribute space for children.

References

Abdulkadiroğlu, A., N. Agarwal, and P. A. Pathak (2017): "The Welfare Effects of Coordinated Assignment: Evidence from the NYC High School Match," *American Economic Review*, 107, 3635–3689.

Abdulkadiroğlu, A., P. A. Pathak, A. E. Roth, and T. Sönmez (2005): "The Boston Public School Match," *American Economic Review Papers and Proceedings*, 95, 368–371.

——— (2011): "Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 101, 195–201.

Autor, D. H., M. M. Kleiner, and A. B. Krueger (2006): "Licensing of Occupations and the Labor Market," *American Economic Review Papers and Proceedings*, 96, 165–169.

——— (2010): "Do Licensing Regulations Reduce Economic Mobility? Evidence from the United States," *British Journal of Industrial Relations*, 48, 676–695.

Baron, J., R. Lombardo, J. P. Ryan, J. Suh, and Q. Valenzuela-Stookey (2024): "Mechanism Reform: An Application to Child Welfare," *Working Paper*.

Becker, G. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–846.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 128, 967–1015.

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

Chade, H. and L. Smith (2006): "Simultaneous Search," *Econometrica*, 74, 1293–1307.

Cox, M., D. Cherry, and J. Orme (2011): "Measuring the willingness to foster children with emotional and behavioral problems," *Children and Youth Services Review*, 33, 59–65.

Damiano, E. and H. Li (2007): "Price Discrimination and Efficient Matching," *Economic Theory*, 30, 243–263.

DeVooght, K. and D. Blazey (2013): "Family Foster Care Reimbursement Rate in the US: A Report from a 2012 National Survey on Family Foster Care Provider Classifications and Rates," *Child Trends Publications.*

Dierks, L., V. Slaugh, and M. U. Ünver (2024): "Child Welfare Platform Design to Improve Outcomes for Children with Disabilities," *Working Paper*.

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

Gibbs, D. and J. Wildfire (2007): "Length of service for foster parents: Using administrative data to understand retention," *Children and Youth Services Review*, 29, 588–599.

Hatfield, J. W. and P. Milgrom (2005): "Matching with Contracts," *American Economic Review*, 95, 913–935.

HIGHSMITH, T. (2024): "Dynamic Envy-Free Permanency in Child Welfare Systems," *Working Paper*.

KELSO, A. S. AND V. P. CRAWFORD (1982): "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 50, 1483–1504.

MACDONALD, D. (2024): "Foster Care: A Dynamic Matching Approach," *Working Paper*.

MENZIO, G. AND S. SHI (2010a): "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, 145, 1453–1494.

——— (2010b): "Directed Search on the Job, Heterogeneity, and Aggregate Fluctuations," *The American Economic Review*, 100, 327–332.

MOEN, E. (1997): "Competitive Search Equilibrium," *Journal of Political Economy*.

MONTGOMERY, J. D. (1991): "Equilibrium Wage Distributions and the Labor Market Search," *Econometrica*, 59, 1719–1744.

O'HARE, W. (2008): "Data on Children in Foster Care from the Census Bureau," *Annie E. Casey Foundation*.

PETERS, M. (1997): "A Revelation Principle for Competing Mechanisms," *Journal of Economic Theory*, 75, 141–168.

ROBINSON-CORTÉS, A. (2021): "Who Gets Placed Where and Why? An Empirical Framework for Foster Care Placement," *Working Paper*.

ROSINSKY, K., M. FISCHER, AND M. HAAS (2023): "Child Welfare Financing SFY 2020: A Survey of Federal, State, and Local Expenditures," *Child Trend Publications*.

SHI, S. (2001): "Frictional Assignment. I. Efficiency," *Journal of Economic Theory*, 98, 232–260.

SHIMER, R. AND L. SMITH (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369.

SLAUGH, V., M. AKAN, O. KESTEN, AND U. UNVER (2016): "The Pennsylvania Adoption Exchange Improves Its Matching Process," *Interfaces*, 46, 133–153.

SMITH, L. (2006): "The Marriage Model with Search Frictions," *Journal of Political Economy*, 114, 1124–1144.

A Appendix: Foster Care in the U.S.

A.1 Overview

During 2020 Federal Fiscal Year (FFY),[31] child welfare agencies across the United States received more than 3.9 million allegations of suspected child abuse or neglect (equivalent to approximately 7.1 million children). Out of these children, 9 percent were removed from their homes and placed into foster care. According to Rosinsky et al. (2023), the national spending on child welfare in 2020 FFY was approximately US$34.1, out of which US$15.2 billion was federally funded, and the remaining was financed directly by States. Furthermore, 45 percent of the national spending was destined to foster care placement expenditure, including payments to foster parents.

Using the Foster Care Files from AFCARS,[32] we observed that in 2020 FFY there were 631,254 children in foster care. On average, these children were almost 7 years old, 49 percent were females, 69 percent were white, and 24 percent were clinically diagnosed with a disability.[33] Thus, based on the disability variable, we can infer that at least 24 percent of children in the U.S. foster care are special needs.[34] During their stay in foster care, 77 percent of these children were placed with foster parents, 9 percent were placed in institutional care, and the remaining had other arrangements. Foster parents caring for children with and without a disability received an average payment of US$1,423 and US$ 2,704 per month, respectively. In this data set, foster parents are not identifiable; only family structure, race and year of birth are reported. Thus, since we do not know how many times a foster parent might appear, we can not provide reliable statistics.

Most of the information regarding foster parents comes from Census data and surveys. Using Census data from 2000, O'Hare (2008) finds that households with foster children, compared to all other households with children, are: less likely to be married-couples, less likely to have a member who finished college, less likely to work full-time, more likely to be low income families, and more likely to receive public assistance in-

[31] October 1, 2019 to September 30, 2020.

[32] AFCARS is a federally mandated data collection system. All fifty US states and the District of Columbia are required to collect data on all children in foster care and all children adopted from foster care.

[33] A disability includes conditions such as blindness, glaucoma, arthritis, multiple sclerosis, down syndrome, personality disorder, attention deficit, and anxiety disorder, among others.

[34] In the majority of the cases, once a child enters the foster care system, a mandatory medical evaluation is performed, therefore we assume that the level of care needed is common knowledge.

come. Now, after conducting a survey of 297 foster mothers, Cox et al. (2011) finds that the average age is 44.1 years old, 88.2 percent are European-American, 75.1 percent are married, 28.9 percent have a bachelor's degree, 33 percent works full-time, and 50.1 percent have an annual family income less than US$50,000.

A.2 Matching Process

Foster care is overseen and managed at the state level by Child Protective Services (CPS). Upon receiving an allegation regarding a child's well-being, CPS assigns a social worker to the case, starting an investigation. If sufficient evidence supporting an accusation is identified, the case is presented to a juvenile or family court. The judge then determines whether the child should be removed from their birth family home and placed in foster care.

In many states, decisions regarding the placement of children are made by social workers. Acting on behalf of the child, the social worker (a) searches for and contacts foster parents, (b) facilitates a meeting between the foster parent and child to assess compatibility, and (c) decides on the placement of the child. In this search process, the social worker can only consider fosters parent who are certified, through a license, to provide care for the child.

Foster parents must obtain a license to provide care for children. The licensing process involves a home study and mandatory training. The home study ensures the foster parent's residence is clean, in good condition, and free from hazards. Initial training, ranging from 15 to 30 hours, covers topics such as agency policies, foster parent roles and responsibilities, and behavior management. The menu of licenses varies across states (for more details see DeVooght and Blazey (2013)). As we mentioned in the introduction, children are grouped by the level of care needed, and transfers vary across groups. These transfers follow the principle that foster parents caring for children with high-needs receive greater transfers.

B Appendix: Analysis of Complete Information

In this section, we prove the results for the complete information case. For each parent y_k with $k = \{1, 2\}$, the designer offers a licenses (λ^k, τ^k). The designer solves the following problem:

$$\max_{\left\{\left(\lambda^k(x_i),\tau^k(x_i)\right)_{i=1}^2\right\}_{k=1}^2} \left\{\sum_{i=1}^{2} \pi^c(\theta_i)\ \frac{\sum_{k=1}^2 \left[u(x_i,y_k)-\tau^k(x_i)\right]\lambda^k(x_i)g(y_k)}{\sum_{k=1}^2 \lambda^k(x_i)\ g(y_k)}\ f(x_i)\right\}$$

subject to [FC]s, [MT]s, and [PC]s defined in the optimization problem 1. Recall that $\pi^p(\theta) = {}^{\pi^c(\theta)}/_{\theta}$. Thus, the objective function can be written as:

$$\max_{\left\{\left(\lambda^k(x_i),\tau^k(x_i)\right)_{i=1}^2\right\}_{k=1}^2} \left\{\sum_{i=1}^{2} \pi^p(\theta_i)\ \sum_{k=1}^{2} \left[u(x_i,y_k)-\tau^k(x_i)\right]\lambda^k(x_i)g(y_k)\right\}$$

Notice from the [PC]s that the expected total transfer must be equal to the expected total cost for each license: $\sum_{i=1}^2 \tau^k(x_i)\lambda^k(x_i)\pi^p(\theta_i) = \sum_{i=1}^2 c(x_i,y_k)\lambda^k(x_i)\pi^p(\theta_i)$. Thus replacing this into the *rearranged* objective function reduces the designer's problem to the following:

$$\max_{\left\{\lambda^k(x_1),\lambda^k(x_2)\right\}_{k=1}^2} \left\{\sum_{i=1}^{2} \pi^p(\theta_i)\ \sum_{k=1}^{2} \left[u(x_i,y_k)-c(x_i,y_k)\right]\lambda^k(x_i)g(y_k)\right\} \quad \text{s.t. [FC]s and [MT]s.}$$

Corollary B.1. *In the first best, the randomization device $\left\{\lambda^k(x_1),\lambda^k(x_2)\right\}_{k=1}^2$ is independent of whether we consider interim or ex-post participation constraints.*

B.1 Proof of Lemma 1

For each (x,k), let $\lambda^k(x)$ be an arbitrary-feasible interior probability that generates a total welfare equal to:

$$W\left(\lambda^1(x_1),\lambda^2(x_1)\right) = \pi^p(\theta_1)\cdot\Big[g(y_1)\ \lambda^1(x_1)\ S(x_1,y_1) + \big(1-g(y_1)\big)\ \lambda^2(x_1)\ S(x_1,y_2)\Big]$$
$$+\ \pi^p(\theta_2)\cdot\Big[g(y_1)\ \big(1-\lambda^1(x_1)\big)\ S(x_2,y_1) + \big(1-g(y_1)\big)\ \big(1-\lambda^2(x_1)\big)\ S(x_2,y_2)\Big],\ \text{where}$$

$$\theta_1 = \frac{g(y_1)\ \lambda^1(x_1) + \big(1-g(y_1)\big)\ \lambda^2(x_1)}{f(x_1)} \quad \text{and} \quad \theta_2 = \frac{g(y_1)\ \big(1-\lambda^1(x_1)\big) + \big(1-g(y_1)\big)\big(1-\lambda^2(x_1)\big)}{1-f(x_1)} \qquad \text{(B.1)}$$

After trembling $\lambda^1(x_1)$ by ε_1 and $\lambda^2(x_1)$ by ε_2 such that $\varepsilon_2 \equiv -{}^{\varepsilon_1 g(y_1)}/_{1-g(y_1)}$, ensuring that the market tightness in each market remains constant, the new total welfare is:

$$W\left(\lambda^1(x_1)+\varepsilon_1,\lambda^2(x_1)+\varepsilon_2\right) = \pi^p(\theta_1)\cdot\Big[g(y_1)\ \lambda^1(x_1)\ S(x_1,y_1) + \big(1-g(y_1)\big)\ \lambda^2(x_1)\ S(x_1,y_2)\Big]$$
$$+\ \pi^p(\theta_2)\cdot\Big[g(y_1)\ \big(1-\lambda^1(x_1)\big)\ S(x_2,y_1) + \big(1-g(y_1)\big)\ \big(1-\lambda^2(x_1)\big)\ S(x_2,y_2)\Big]$$
$$+\ \varepsilon_1\ g(y_1)\ \Big(\pi^p(\theta_2)\ \big[S(x_2,y_2)-S(x_2,y_1)\big] - \pi^p(\theta_1)\ \big[S(x_1,y_2)-S(x_1,y_1)\big]\Big)$$

Thus, the change in welfare is equal to:

$$\Delta_W = W\big(\lambda^1(x_1)+\varepsilon_1, \lambda^2(x_1)+\varepsilon_2\big) - W\big(\lambda^1(x_1), \lambda^2(x_1)\big)$$
$$= \varepsilon_1\, g(y_1)\, \underbrace{\Big(\pi^p(\theta_2)\, \big[S(x_2,y_2) - S(x_2,y_1)\big] - \pi^p(\theta_1)\, \big[S(x_1,y_2) - S(x_1,y_1)\big]\Big)}_{Z^{CI}(\theta_1)}$$

where θ_1 and θ_2 are defined as in Equation B.1. Note that $\theta_2 = {}^{1-f(x_1)\theta_1}/_{1-f(x_1)}$, thus Z^{CI} can be written as a function of only θ_1. It is easy to see that $Z^{CI}(\theta_1)$ is strictly increasing in θ_1. Therefore, $Z^{CI}(\theta_1^{\max}) \geq Z^{CI}(\theta_1) \geq Z^{CI}(0)$ for any $\theta_1 \in [0, \theta_1^{\max}]$ where $\theta_1^{\max} = {}^{1}/_{f(x_1)}$. Now, we analyze three cases:

1. Suppose $Z^{CI}(\theta_1) > 0$. Then, pick $\varepsilon_1 > 0$ with $\varepsilon_2 = {}^{-\varepsilon_1 g(y_1)}/_{1-g(y_1)}$ such that either $\hat{\lambda}^1(x_1) \equiv \lambda^1(x_1) + \varepsilon_1 = 1$ or $\hat{\lambda}^2(x_1) \equiv \lambda^2(x_1) + \varepsilon_2 = 0$. In the former case, $\hat{\lambda}^1(x_2) = 0$ and $\hat{\lambda}^2(x_2) \in (0,1)$; and in the latter case, $\hat{\lambda}^1(x_2) \in (0,1)$ and $\hat{\lambda}^2(x_2) = 1$. In both cases, the definition of PAM is satisfied.

2. Suppose $Z^{CI}(\theta_1) < 0$. Then, pick $\varepsilon_1 < 0$ with $\varepsilon_2 = {}^{-\varepsilon_1 g(y_1)}/_{1-g(y_1)}$ such that either $\hat{\lambda}^1(x_1) \equiv \lambda^1(x_1) + \varepsilon_1 = 0$ or $\hat{\lambda}^2(x_1) \equiv \lambda^2(x_1) + \varepsilon_2 = 1$. In the former case, $\hat{\lambda}^1(x_2) = 1$ and $\hat{\lambda}^2(x_2) \in (0,1)$; and in the latter case, $\hat{\lambda}^1(x_2) \in (0,1)$ and $\hat{\lambda}^2(x_2) = 0$. In both cases, the definition of NAM is satisfied.

3. Suppose $Z^{CI}(\theta_1) = 0$. We show that an interior randomization device can not be an equilibrium. To see this, first tremble $\lambda^1(x_1)$ by ε_1, and calculate welfare:

$$W\big(\lambda^1(x_1)+\varepsilon_1, \lambda^2(x_1)\big) = \pi^p(\hat{\theta}_1)\cdot\Big[g(y_1)\,\lambda^1(x_1)\,S(x_1,y_1) + \big(1-g(y_1)\big)\,\lambda^2(x_1)\,S(x_1,y_2)\Big]$$
$$+\,\pi^p(\hat{\theta}_2)\cdot\Big[g(y_1)\,\big(1-\lambda^1(x_1)\big)\,S(x_2,y_1) + \big(1-g(y_1)\big)\,\big(1-\lambda^2(x_1)\big)\,S(x_2,y_2)\Big]$$
$$+\,\varepsilon_1 g(y_1)\Big[\pi^p(\hat{\theta}_1)\,S(x_1,y_1) - \pi^p(\hat{\theta}_2)\,S(x_2,y_1)\Big]$$

where $\hat{\theta}_1 = \theta_1 + {}^{\varepsilon_1 g(y_1)}/_{f(x_1)}$, $\hat{\theta}_2 = \theta_2 - {}^{\varepsilon_1 g(y_1)}/_{1-f(x_1)}$, and θ_1, θ_2 are defined as in Equation B.1. Now, let's tremble $\lambda^2(x_1)$ by ε_2, and calculate welfare:

$$W(\lambda^2(x_1), \lambda^2(x_1)+\varepsilon_2) = \pi^p(\tilde{\theta}_1)\cdot\Big[g(y_1)\,\lambda^1(x_1)\,S(x_1,y_1) + \big(1-g(y_1)\big)\,\lambda^2(x_1)\,S(x_1,y_2)\Big]$$
$$+\,\pi^p(\tilde{\theta}_2)\cdot\Big[g(y_1)\,\big(1-\lambda^1(x_1)\big)\,S(x_2,y_1) + \big(1-g(y_1)\big)\,\big(1-\lambda^2(x_1)\big)\,S(x_2,y_2)\Big]$$
$$+\,\varepsilon_2\big(1-g(y_1)\big)\Big[\pi^p(\tilde{\theta}_1)\,S(x_1,y_2) - \pi^p(\tilde{\theta}_2)\,S(x_2,y_2)\Big]$$

where $\tilde{\theta}_1 = \theta_1 + {}^{\varepsilon_2(1-g(y_1))}/_{f(x_1)}$, $\tilde{\theta}_2 = \theta_2 - {}^{\varepsilon_2(1-g(y_1))}/_{1-f(x_1)}$, and θ_1, θ_2 are defined as in Equation B.1. For any small ε_1 with $\varepsilon_2 \equiv {}^{\varepsilon_1 g(y_1)}/_{1-g(y_1)}$, it follows that $\hat{\theta}_1 = \tilde{\theta}_1$ and $\hat{\theta}_2 = \tilde{\theta}_2$. Pick such ε_2. Then, increasing $\lambda^1(x_1)$ is marginally more profitable than increasing

$\lambda^2(x_1)$ if and only if

$$\underbrace{\pi^p(\hat{\theta}_2)\cdot\big[S(x_2,y_2)-S(x_2,y_1)\big]-\pi^p(\hat{\theta}_1)\cdot\big[S(x_1,y_2)-S(x_1,y_1)\big]}_{Z^{CI}(\hat{\theta}_1)}\geq 0$$

Since $Z^{CI}(\hat{\theta}_1) > Z^{CI}(\theta_1) = 0$, then the inequality holds. Therefore, at least one of the partial derivatives of W at $\big(\lambda^1(x_1), \lambda^2(x_1)\big)$ is non-zero, meaning that $\big(\lambda^1(x_1), \lambda^2(x_1)\big)$ at $Z^{CI}(\theta_1) = 0$ is not an equilibrium. This finishes the proof.

B.2 Proof of Proposition 1

By assumption $S(x, y)$ is increasing in y, thus $Z^{CI}(\theta_1)$ is increasing in θ_1. Therefore, items (i) to (iii) from the previous proof of Lemma 1 apply here.

B.3 Proof of Corollary 1

Notice that, $Z^{CI}(\theta_1)$ is increasing in θ_1 reaching its minimum value at $\theta_1 = 0$, and when $\theta_1 = 0$ it follows that $\pi^p(0) = 1$ and $\theta_2 = {}^{1}/_{1-f(x_1)}$. Therefore, from Proposition 1, we can ensure PAM by imposing that the following inequality must hold:

$$\pi^p\big({}^{1}/_{1-f(x_1)}\big)\cdot\big[S(x_2,y_2)-S(x_2,y_1)\big]-\big[S(x_1,y_2)-S(x_1,y_1)\big]\geq 0$$

Notice, $Z^{CI}(\theta_1)$ reaches its maximum value at $\theta_1 = {}^{1}/_{f(x_1)}$, and thus the condition for NAM also simply follows.

B.4 Proof of Proposition 2

The designer solves the following problem:

$$\max_{\left\{\left(\lambda^k(x_i),\tau^k(x_i)\right)_{i=1}^2\right\}_{k=1}^2}\left\{\sum_{i=1}^{2}\pi^c(\theta_i)\,\frac{\sum_{k=1}^{2}\left[u(x_i,y_k)-\tau^k(x_i)\right]\lambda^k(x_i)g(y_k)}{\sum_{k=1}^{2}\lambda^k(x_i)\,g(y_k)}\,f(x_i)\right\}$$

subject to [FC], [MT], and [PC]. We show that at the optimum, the participation constraints hold with equality. By contradiction, suppose that for some license k, the [PC] holds with strict inequality: $\sum_{i=1}^2 \tau^k(x_i)\lambda^k(x_i)\pi^p(\theta_i) > \sum_{i=1}^2 c(x_i, y_k)\lambda^k(x_i)\pi^p(\theta_i)$ Then, the designer can decrease $\tau^k(x_1)$ and $\tau^k(x_2)$ by a small $\varepsilon > 0$ satisfying the constraint while increasing the objective function. A contradiction. Therefore, the optimal transfers can be pinned-down by the [PC] which hold with equality.

C Appendix: Analysis of Private Information

First, it is useful to understand who has incentives to mimic whom under the first best menu of licenses. Recall the incentive compatibility constraint [IC] for $k \neq k' = 1, 2$: $\sum_{i=1}^{2} \left[\tau^k(x_i) - c(x_i, y_k)\right] \lambda^k(x_i)\pi^p(\theta_i) \geq \sum_{i=1}^{2} \left[\tau^{k'}(x_i) - c(x_i, y_k)\right] \lambda^{k'}(x_i)\pi^p(\theta_i)$ and the participation constraint [PC] for $k = 1, 2$: $\sum_{i=1}^{2} \left[\tau^k(x_i) - c(x_i, y_k)\right] \lambda^k(x_i)\pi^p(\theta_i) \geq 0$.

In the complete information case, [PC]s holds with equality. Now, plugging [PC](k) and [PC](k') into [IC](k) yields the following inequality:

$$0 \geq \Big[c(x_1, y_{k'}) - c(x_1, y_k)\Big] \lambda^{k'}(x_1)\pi^p(\theta_1) + \Big[c(x_2, y_{k'}) - c(x_2, y_k)\Big] \lambda^{k'}(x_2)\pi^p(\theta_2)$$

Since $c(x, y)$ is decreasing in y, the inequality holds for $k = 1$ but not for $k = 2$. Thus, under the first best, type-y_2 parents have incentives to mimic type-y_1 parents. Notice the [IC] for high-ability and the [PC] for low-ability parents hold with equality in equilibrium (see Proof of Proposition 4). Thus, plugging objects from [PC1] into [IC2]:

$$\tau^2(x_1)\lambda^2(x_1)\pi^p(\theta_1) + \tau^2(x_2)\lambda^2(x_2)\pi^p(\theta_2) = c(x_1, y_2)\lambda^2(x_1)\pi^p(\theta_1) + c(x_2, y_2)\lambda^2(x_2)\pi^p(\theta_2) + \Big[c(x_1, y_1) - c(x_1, y_2)\Big] \lambda^1(x_1)\pi^p(\theta_1) + \Big[c(x_2, y_1) - c(x_2, y_2)\Big] \lambda^1(x_2)\pi^p(\theta_2)$$

Now, replacing the restrictions into the objective function, the designer solves:

$$\max_{\{\lambda^k(x_1), \lambda^k(x_2)\}_{k=1}^2} \Bigg\{ \sum_{i=1}^{2} \pi^p(\theta_i) \left[\sum_{k=1}^{2} \Big(\underbrace{u(x_i, y_k) - c(x_i, y_k)}_{S(x,y)} \Big) \lambda^k(x_i)\ g(y_k) \right] - \Big[c(x_1, y_1) - c(x_1, y_2)\Big] \lambda^1(x_1)\pi^p(\theta_1) g(y_2) - \Big[c(x_2, y_1) - c(x_2, y_2)\Big] \lambda^1(x_2)\pi^p(\theta_2) g(y_2) \Bigg\}$$

subject to [FC], [MT], and some additional constraints [AC] where

$$\frac{c(x_2, y_2) - c(x_2, y_1)}{c(x_1, y_2) - c(x_1, y_1)} \geq \frac{1}{\pi^p\left(1/f(x_2)\right)} \text{ if } \lambda^2(x_2) > \lambda^1(x_2) \text{ and } \frac{c(x_1, y_2) - c(x_1, y_1)}{c(x_2, y_2) - c(x_2, y_1)} \geq \frac{1}{\pi^p\left(1/f(x_1)\right)} \text{ if } \lambda^2(x_2) < \lambda^1(x_2) \quad \text{(C.1)}$$

This additional constraint [AC] ensures that the [IC] for low-ability parents is satisfied when the [IC] for high ability parents holds (see Proof of Proposition 4). Since the objective function is independent of the transfers after incorporating the participation constraints, the following is immediate:

Corollary C.1. *With private information, the randomization device* $\left\{\lambda^k(x_1), \lambda^k(x_2)\right\}_{k=1}^2$ *is independent of whether interim or ex-post participation constraints are implemented.*

C.1 Proof of Lemma 1 under Private Information

We can establish Lemma 1 for the private information case.

Lemma C.1. *In the private information setting, for at least one of the licenses, the optimal randomization rule yields a corner solution.*

For each (x, k), let $\lambda^k(x_1) \in (0,1)$ be an arbitrary-feasible interior probability that generates a total welfare equal to:

$$\hat{W}\big(\lambda^1(x_1), \lambda^2(x_1)\big) = \pi^p(\theta_1)\cdot\Big[g(y_1)\,\lambda^1(x_1)\,S(x_1,y_1) + \big(1-g(y_1)\big)\,\lambda^2(x_1)\,S(x_1,y_2)\Big]$$
$$+\,\pi^p(\theta_2)\cdot\Big[g(y_1)\,\big(1-\lambda^1(x_1)\big)\,S(x_2,y_1) + \big(1-g(y_1)\big)\,\big(1-\lambda^2(x_1)\big)\,S(x_2,y_2)\Big]$$
$$-\Big[c(x_1,y_1)-c(x_1,y_2)\Big]\lambda^1(x_1)\pi^p(\theta_1)g(y_2) - \Big[c(x_2,y_1)-c(x_2,y_2)\Big]\big(1-\lambda^1(x_1)\big)\pi^p(\theta_2)g(y_2)$$

where $\theta_1 = \frac{g(y_1)\,\lambda^1(x_1)+\big(1-g(y_1)\big)\,\lambda^2(x_1)}{f(x_1)}$, and $\theta_2 = \frac{g(y_1)\,\big(1-\lambda^1(x_1)\big)+\big(1-g(y_1)\big)\big(1-\lambda^2(x_1)\big)}{1-f(x_1)}$. As in the complete information, we tremble $\lambda^1(x_1)$ by ε_1 and $\lambda^2(x_1)$ by ε_2 such that $\varepsilon_2 \equiv {-\varepsilon_1 g(y_1)}/{1-g(y_1)}$ ensuring that the market tightness in each submarket remains constant. The new total welfare is:

$$\hat{W}\big(\lambda^1(x_1)+\varepsilon_1, \lambda^2(x_1)+\varepsilon_2\big) = \pi^p(\theta_1)\cdot\Big[g(y_1)\,\lambda^1(x_1)\,S(x_1,y_1) + \big(1-g(y_1)\big)\,\lambda^2(x_1)\,S(x_1,y_2)\Big]$$
$$+\,\pi^p(\theta_2)\cdot\Big[g(y_1)\,\big(1-\lambda^1(x_1)\big)\,S(x_2,y_1) + \big(1-g(y_1)\big)\,\big(1-\lambda^2(x_1)\big)\,S(x_2,y_2)\Big]$$
$$-\Big[c(x_1,y_1)-c(x_1,y_2)\Big]\lambda^1(x_1)\pi^p(\theta_1)g(y_2) - \Big[c(x_2,y_1)-c(x_2,y_2)\Big]\big(1-\lambda^1(x_1)\big)\pi^p(\theta_2)g(y_2)$$
$$+\,\varepsilon_1 g(y_1)\Bigg\{\pi^p(\theta_2)\Big[S(x_2,y_2)-S(x_2,y_1)+\frac{g(y_2)}{g(y_1)}\Big(c(x_2,y_1)-c(x_2,y_2)\Big)\Big]$$
$$-\,\pi^p(\theta_1)\Big[S(x_1,y_2)-S(x_1,y_1)+\frac{g(y_2)}{g(y_1)}\Big(c(x_1,y_1)-c(x_1,y_2)\Big)\Big]\Bigg\}$$

Thus, the change in welfare is equal to:

$$\Delta_{\hat{W}} = \varepsilon_1 g(y_1)\Bigg\{\underbrace{\pi^p(\theta_2)\Big[S(x_2,y_2)-S(x_2,y_1)+\frac{g(y_2)}{g(y_1)}\Big(c(x_2,y_1)-c(x_2,y_2)\Big)\Big] - \pi^p(\theta_1)\Big[S(x_1,y_2)-S(x_1,y_1)+\frac{g(y_2)}{g(y_1)}\Big(c(x_1,y_1)-c(x_1,y_2)\Big)\Big]}_{Z^{PI}(\theta_1)}\Bigg\}$$

where θ_1 and θ_2 are defined as above. Note that $Z^{PI}(\theta_1)$ is strictly increasing in θ_1. Therefore, $Z^{PI}(\theta_1^{\max}) \geq Z^{PI}(\theta_1) \geq Z^{PI}(0)$ for any $\theta_1 \in [0, \theta_1^{\max}]$ where $\theta_1^{\max} = {1}/{f(x_1)}$.

Now, we analyze three cases:

1. Suppose $Z^{PI}(\theta_1) > 0$. Then, pick $\varepsilon_1 > 0$ with $\varepsilon_2 \equiv -\varepsilon_1 g(y_1)/1-g(y_1)$ such that either $\hat{\lambda}^1(x_1) \equiv \lambda^1(x_1) + \varepsilon_1 = 1$ or $\hat{\lambda}^2(x_1) \equiv \lambda^2(x_1) + \varepsilon_2 = 0$. In the former case, $\hat{\lambda}^1(x_2) = 0$ and $\hat{\lambda}^2(x_2) \in (0,1)$; and in the latter case, $\hat{\lambda}^1(x_2) \in (0,1)$ and $\hat{\lambda}^2(x_2) = 1$. In both cases, the definition of PAM is satisfied.

2. Suppose $Z^{PI}(\theta_1) < 0$. Then, pick $\varepsilon_1 < 0$ with $\varepsilon_2 \equiv -\varepsilon_1 g(y_1)/1-g(y_1)$ such that either $\hat{\lambda}^1(x_1) \equiv \lambda^1(x_1) + \varepsilon_1 = 0$ or $\hat{\lambda}^2(x_1) \equiv \lambda^2(x_1) + \varepsilon_2 = 1$. In the former case, $\hat{\lambda}^1(x_2) = 1$ and $\hat{\lambda}^2(x_2) \in (0,1)$; and in the latter case, $\hat{\lambda}^1(x_2) \in (0,1)$ and $\hat{\lambda}^2(x_2) = 0$. In both cases, the definition of NAM is satisfied.

3. Suppose $Z^{PI}(\theta) = 0$. We show that an interior randomization device can not be an equilibrium. To see this, first tremble $\lambda^1(x_1)$ by ε_1, and calculate welfare:

$$\begin{aligned}
\hat{W}\big(\lambda^1(x_1)+\varepsilon_1, \lambda^2(x_1)\big) &= \pi^p(\hat{\theta}_1)\cdot\Big[g(y_1)\,\lambda^1(x_1)\,S(x_1,y_1) + \big(1-g(y_1)\big)\,\lambda^2(x_1)\,S(x_1,y_2)\Big] \\
&\quad + \pi^p(\hat{\theta}_2)\cdot\Big[g(y_1)\,\big(1-\lambda^1(x_1)\big)\,S(x_2,y_1) + \big(1-g(y_1)\big)\,\big(1-\lambda^2(x_1)\big)\,S(x_2,y_2)\Big] \\
&\quad - \Big[c(x_1,y_1)-c(x_1,y_2)\Big]\lambda^1(x_1)\pi^p(\hat{\theta}_1)g(y_2) - \Big[c(x_2,y_1)-c(x_2,y_2)\Big]\big(1-\lambda^1(x_1)\big)\pi^p(\hat{\theta}_2)g(y_2) \\
&\quad + \varepsilon_1 g(y_1)\Big[\pi^p(\hat{\theta}_1)\,S(x_1,y_1) - \pi^p(\hat{\theta}_2)\,S(x_2,y_1)\Big] \\
&\quad + \varepsilon_1 g(y_2)\Big\{\pi^p(\hat{\theta}_2)\Big[c(x_2,y_1)-c(x_2,y_2)\Big] - \pi^p(\hat{\theta}_1)\Big[c(x_1,y_1)-c(x_1,y_2)\Big]\Big\}
\end{aligned}$$

where $\hat{\theta}_1 = \theta_1 + \varepsilon_1 g(y_1)/f(x_1)$, $\hat{\theta}_2 = \theta_2 - \varepsilon_1 g(y_1)/1-f(x_1)$, and θ_1, θ_2 are defined as above. Now, let's tremble $\lambda^2(x_1)$ by ε_2, and calculate welfare:

$$\begin{aligned}
\hat{W}(\lambda^2(x_1), \lambda^2(x_1)+\varepsilon_2) &= \pi^p(\tilde{\theta}_1)\cdot\Big[g(y_1)\,\lambda^1(x_1)\,S(x_1,y_1) + \big(1-g(y_1)\big)\,\lambda^2(x_1)\,S(x_1,y_2)\Big] \\
&\quad + \pi^p(\tilde{\theta}_2)\cdot\Big[g(y_1)\,\big(1-\lambda^1(x_1)\big)\,S(x_2,y_1) + \big(1-g(y_1)\big)\,\big(1-\lambda^2(x_1)\big)\,S(x_2,y_2)\Big] \\
&\quad - \Big[c(x_1,y_1)-c(x_1,y_2)\Big]\lambda^1(x_1)\pi^p(\tilde{\theta}_1)g(y_2) - \Big[c(x_2,y_1)-c(x_2,y_2)\Big]\big(1-\lambda^1(x_1)\big)\pi^p(\tilde{\theta}_2)g(y_2) \\
&\quad + \varepsilon_2\big(1-g(y_1)\big)\Big[\pi^p(\tilde{\theta}_1)\,S(x_1,y_2) - \pi^p(\tilde{\theta}_2)\,S(x_2,y_2)\Big]
\end{aligned}$$

where $\tilde{\theta}_1 = \theta_1 + \varepsilon_2(1-g(y_1))/f(x_1)$, $\tilde{\theta}_2 = \theta_2 - \varepsilon_2(1-g(y_1))/1-f(x_1)$, and θ_1, θ_2 are defined above. For any small ε_1 with $\varepsilon_2 \equiv \varepsilon_1 g(y_1)/1-g(y_1)$, it follows that $\hat{\theta}_1 = \tilde{\theta}_1$ and $\hat{\theta}_2 = \tilde{\theta}_2$. Pick such ε_2. Then, increasing $\lambda^1(x_1)$ is marginally more profitable than increasing $\lambda^2(x_1)$ if and only if

$$\underbrace{\pi^p(\theta_2)\Big[S(x_2,y_2)-S(x_2,y_1)+\frac{g(y_2)}{g(y_1)}\Big(c(x_2,y_1)-c(x_2,y_2)\Big)\Big] -\pi^p(\theta_1)\Big[S(x_1,y_2)-S(x_1,y_1)+\frac{g(y_2)}{g(y_1)}\Big(c(x_1,y_1)-c(x_1,y_2)\Big)\Big]}_{Z^{PI}(\theta_1)} \geq 0$$

Since $Z^{PI}(\hat{\theta}_1) > Z^{PI}(\theta_1) = 0$, then the inequality holds.[35] Therefore, at least one of the partial derivatives of W at $\left(\lambda^1(x_1), \lambda^2(x_1)\right)$ is non-zero, meaning that $\left(\lambda^1(x_1), \lambda^2(x_1)\right)$ at $Z^{PI}(\theta_1) = 0$ is not an equilibrium. This finishes the proof.

C.2 Proof of Proposition 1 under Private Information

We can establish Proposition 1 for the private information case. Let $\widehat{\theta}_1$ be such that $Z^{PI}(\widehat{\theta}_1) = 0$, then the following result holds:

Proposition C.1. *In the private information setting, let θ_1^{**} be the equilibrium market tightness.* ***(i)*** *If $\theta_1^{**} > \widehat{\theta}_1$ then the equilibrium sorting exhibits PAM.* ***(ii)*** *If $\theta_1^{**} < \widehat{\theta}_1$ then the equilibrium sorting exhibits NAM.* ***(iii)*** *$\theta_1^{**} = \widehat{\theta}_1$ is never optimal.*

Recall $S(x, y)$ is increasing in y, so is $Z^{PI}(\theta_1)$ simply implying Proposition C.1.

C.3 Proof of Corollary 2

$Z^{PI}(\theta_1)$ is increasing in θ_1 reaching its minimum value at $\theta_1 = 0$, and when $\theta_1 = 0$ it follows that $\pi^p(0) = 1$ and $\theta_2 = {}^1/_{1-f(x_1)}$. Therefore, from Proposition C.1, we can ensure PAM by imposing that the following inequality must hold:

$$\pi^p\left({}^1/_{f(x_2)}\right)\Big[S(x_2,y_2)-S(x_2,y_1)+{}^{g(y_2)}/_{g(y_1)}\Big(c(x_2,y_1)-c(x_2,y_2)\Big)\Big] - \Big[S(x_1,y_2)-S(x_1,y_1)+{}^{g(y_2)}/_{g(y_1)}\Big(c(x_1,y_1)-c(x_1,y_2)\Big)\Big] \geq 0$$

Now, $Z^{PI}(\theta_1)$ reaches its maximum value at $\theta_1 = {}^1/_{f(x_1)}$.Therefore, from Proposition C.1, we can ensure NAM by imposing that the following inequality must hold:

$$\Big[S(x_2,y_2)-S(x_2,y_1)+{}^{g(y_2)}/_{g(y_1)}\Big(c(x_2,y_1)-c(x_2,y_2)\Big)\Big] - \pi^p\left({}^1/_{f(x_1)}\right)\Big[S(x_1,y_2)-S(x_1,y_1)+{}^{g(y_2)}/_{g(y_1)}\Big(c(x_1,y_1)-c(x_1,y_2)\Big)\Big] \leq 0$$

[35] It is easy to verify the equivalence between the expression here and Equation (4)

C.4 Proof of Proposition 4

The designer solves the following problem:

$$\max_{\left\{ \left(\lambda^k(x_i), \tau^k(x_i)\right)_{i=1}^2 \right\}_{k=1}^2} \left\{ \sum_{i=1}^{2} \pi^c(\theta_i) \, \frac{\sum_{k=1}^{2} \left[u(x_i, y_k) - \tau^k(x_i)\right] \lambda^k(x_i) g(y_k)}{\sum_{k=1}^{2} \lambda^k(x_i) \, g(y_k)} \, f(x_i) \right\}$$

subject to [FC], [MT],[PC], and [IC]. We will analyze the constraints in this maximization problem. First, consider the [IC]s for low- and high-ability parents, respectively:

$$\sum_{i=1}^{2} c(x_i, y_1) \left[\lambda^2(x_i) - \lambda^1(x_i)\right] \pi^p(\theta_i) \geq \sum_{i=1}^{2} \left[\tau^2(x_i) \, \lambda^2(x_i) - \tau^1(x_i) \, \lambda^1(x_i)\right] \pi^p(\theta_i)$$

$$\sum_{i=1}^{2} \left[\tau^2(x_i) \, \lambda^2(x_i) - \tau^1(x_i) \, \lambda^1(x_i)\right] \pi^p(\theta_i) \geq \sum_{i=1}^{2} c(x_i, y_2) \left[\lambda^2(x_i) - \lambda^1(x_i)\right] \pi^p(\theta_i)$$

From these inequalities, we get the following expression, and we continue rearranging:

$$\sum_{i=1}^{2} c(x_i, y_1) \left[\lambda^2(x_i) - \lambda^1(x_i)\right] \pi^p(\theta_i) \geq \sum_{i=1}^{2} c(x_i, y_2) \left[\lambda^2(x_i) - \lambda^1(x_i)\right] \pi^p(\theta_i)$$

$$\Rightarrow c(x_1, y_1) \left[\lambda^2(x_1) - \lambda^1(x_1)\right] \pi^p(\theta_1) + c(x_2, y_1) \left[\lambda^2(x_2) - \lambda^1(x_2)\right] \pi^p(\theta_2) \geq$$
$$c(x_1, y_2) \left[\lambda^2(x_1) - \lambda^1(x_1)\right] \pi^p(\theta_1) + c(x_2, y_2) \left[\lambda^2(x_2) - \lambda^1(x_2)\right] \pi^p(\theta_2)$$

$$\Rightarrow \left[c(x_2, y_1) - c(x_2, y_2)\right] \cdot \left[\lambda^2(x_2) - \lambda^1(x_2)\right] \pi^p(\theta_2) \geq$$
$$\left[c(x_1, y_2) - c(x_1, y_1)\right] \cdot \left[\lambda^2(x_1) - \lambda^1(x_1)\right] \pi^p(\theta_1)$$

Note that $\lambda^2(x_1) - \lambda^1(x_1) = 1 - \lambda^2(x_2) - [1 - \lambda^1(x_2)] = \lambda^1(x_2) - \lambda^2(x_2)$, hence replacing in the previous inequality yields:

$$\left[c(x_2, y_1) - c(x_2, y_2)\right] \cdot \left[\lambda^2(x_2) - \lambda^1(x_2)\right] \pi^p(\theta_2) \geq$$
$$\left[c(x_1, y_1) - c(x_1, y_2)\right] \cdot \left[\lambda^2(x_2) - \lambda^1(x_2)\right] \pi^p(\theta_1) \quad \text{(C.2)}$$

This inequality depends on the sign of the term $[\lambda^2(x_2) - \lambda^1(x_2)]$, which defines PAM and NAM. Hence, consider the following cases:

Case 1. Suppose $\lambda^2(x_2) - \lambda^1(x_2)$ is positive. Then, Equation C.2 reduces to: $\left[c(x_2, y_1) - c(x_2, y_2)\right] \cdot \pi^p(\theta_2) \geq \left[c(x_1, y_1) - c(x_1, y_2)\right] \cdot \pi^p(\theta_1)$ which is satisfied if the following holds:

$$\frac{c(x_2, y_2) - c(x_2, y_1)}{c(x_1, y_2) - c(x_1, y_1)} \geq \frac{1}{\pi^p \left(1/f(x_2)\right)} \quad \text{(C.3)}$$

Case 2. Suppose $\lambda^2(x_2) - \lambda^1(x_2)$ is negative. Then, Equation C.2 reduces to: $\left[c(x_1, y_1) - c(x_1, y_2)\right] \cdot \pi^p(\theta_1) \geq \left[c(x_2, y_1) - c(x_2, y_2)\right] \cdot \pi^p(\theta_2)$ which is satisfied if the following holds:

$$\frac{c(x_1,y_2)-c(x_1,y_1)}{c(x_2,y_2)-c(x_2,y_1)} \geq \frac{1}{\pi^p\left(1/f(x_1)\right)} \tag{C.4}$$

Now, we show that the [PC] for low-ability parents, and the [IC] for high-ability parents imply the [PC] for high-ability parents:

$$\begin{aligned}\sum_{i=1}^{2}\left[\tau^2(x_i)-c(x_i,y_2)\right]\lambda^2(x_i)\pi^p(\theta_i) &\geq \sum_{i=1}^{2}\left[\tau^1(x_i)-c(x_i,y_2)\right]\lambda^1(x_i)\pi^p(\theta_i)\\ &\geq \sum_{i=1}^{2}\left[\tau^1(x_i)-c(x_i,y_1)\right]\lambda^1(x_i)\pi^p(\theta_i) \geq 0\end{aligned}$$

Thus, we can ignore the [PC] for high-ability parents. Next, suppose that the [IC] for high-ability parents holds with strict inequality:

$$\sum_{i=1}^{2}\left[\tau^2(x_i)-c(x_i,y_2)\right]\lambda^2(x_i)\pi^p(\theta_i) > \sum_{i=1}^{2}\left[\tau^1(x_i)-c(x_i,y_2)\right]\lambda^1(x_i)\pi^p(\theta_i)$$

Then, the designer can decrease $\tau^2(x_1)$ and $\tau^2(x_2)$ by a small $\varepsilon > 0$ satisfying the constraint while increasing the objective function. A contradiction. Therefore, the [IC] for high-ability parents holds with equality at the optimum.

Similarly, suppose that the [PC] for low-ability parents holds with strict inequality: $\sum_{i=1}^{2}\left[\tau^1(x_i)-c(x_i,y_1)\right]\lambda^1(x_i)\pi^p(\theta_i) > 0$. Then, the designer can decrease $\tau^1(x_1)$ and $\tau^1(x_2)$ by a small $\varepsilon > 0$ satisfying the constraint while increasing the objective function. A contradiction. Therefore, the [PC] for low-ability parents holds with equality at the optimum.

Lastly, we show that the [IC] for high-ability parents combined with Equations C.3 and C.4 imply the [IC] for low-ability parents. Recall the [IC] for high-ability parents:

$$\begin{aligned}&\sum_{i=1}^{2}\left[\tau^2(x_i)-c(x_i,y_2)\right]\lambda^2(x_i)\pi^p(\theta_i) = \sum_{i=1}^{2}\left[\tau^1(x_i)-c(x_i,y_2)\right]\lambda^1(x_i)\pi^p(\theta_i)\\ &\Rightarrow \sum_{i=1}^{2}\left[\tau^2(x_i)\lambda^2(x_i)-\tau^1(x_i)\lambda^1(x_i)\right]\pi^p(\theta_i) = \sum_{i=1}^{2}c(x_i,y_2)\left[\lambda^2(x_i)-\lambda^1(x_i)\right]\pi^p(\theta_i)\end{aligned}$$

The right-hand side of the previous equation can be written as:

$$\begin{aligned}&c(x_1,y_2)\left[\lambda^2(x_1)-\lambda^1(x_1)\right]\pi^p(\theta_1)+c(x_2,y_2)\left[\lambda^2(x_2)-\lambda^1(x_2)\right]\pi^p(\theta_2)\\ &\quad\Rightarrow c(x_2,y_2)\left[\lambda^2(x_2)-\lambda^1(x_2)\right]\pi^p(\theta_2)-c(x_1,y_2)\left[\lambda^2(x_2)-\lambda^1(x_2)\right]\pi^p(\theta_1)\\ &\qquad\Rightarrow \left[c(x_2,y_2)\pi^p(\theta_2)-c(x_1,y_2)\pi^p(\theta_1)\right]\cdot\left[\lambda^2(x_2)-\lambda^1(x_2)\right]\end{aligned}$$

Thus, the [IC] for high-ability parents can be written as:

$$\sum_{i=1}^{2}\left[\tau^2(x_i)\lambda^2(x_i)-\tau^1(x_i)\lambda^1(x_i)\right]\pi^p(\theta_i)=\Big[c(x_2,y_2)\pi^p(\theta_2)-c(x_1,y_2)\pi^p(\theta_1)\Big]\cdot\left[\lambda^2(x_2)-\lambda^1(x_2)\right]$$

Notice there are two possibilities: Either (i) $\lambda^2(x_2)>\lambda^1(x_2)$, in which case Equation C.3 ensures, after some algebra, that the [IC] for low-ability parents hold; or (ii) $\lambda^2(x_2)<\lambda^1(x_2)$, in which case Equation C.4 ensures, after some algebra, that the [IC] for low-ability parents hold. Therefore, we can drop the [IC] for low-ability parents.

D Appendix: Continuous Type of Parents

First, let's rewrite the constraints the designer faces when solving Equation 6:

$$[\text{FC}]\quad \tau(x,y)\geq 0 \text{ and } \lambda(x,y)\geq 0 \text{ for all } (x,y),\ \text{ and} \sum_{i=1}^{2}\lambda(x_i,y)=1 \text{ for all } y\in Y.$$

$$[\text{MT}]\quad \theta_i=\frac{1}{f(x)}\cdot\int_{\underline{y}}^{\overline{y}}\lambda(x,y)g(y)\,dy\ ,\ \text{ for all } i\in\{1,2\}$$

$$[\text{PC}]\quad \sum_{i=1}^{2}\Big[\tau(x_i,y)-c(x_i,y)\Big]\lambda(x_i,y)\pi^p(\theta_i)\geq 0\ ,\ \text{ for all } y\in Y.$$

D.1 Proof of Lemma 2

Let $\lambda(x,y)$ be an arbitrary interior allocation for any y, that is, $\lambda(x_1,y)\in(0,1)$, and thus $\lambda(x_2,y)=1-\lambda(x_1,y)\in(0,1)$ for any y by [FC]. Define a perturbation function $\varepsilon:Y\to(0,1)$ such that $\int\limits_{y\in Y}\varepsilon(y)g(y)dy=0$.

Consider the allocations $\lambda(y)\equiv\big(\lambda(x_1,y),\lambda(x_2,y)\big)$ and $\tilde{\lambda}(y)\equiv\big(\lambda(x_1,y)-\varepsilon(y),\lambda(x_2,y)+\varepsilon(y)\big)$. Notice, the market tightness derived by allocations $\lambda(y)$ and $\tilde{\lambda}(y)$ is the same: $\theta_i=\frac{1}{f(x_i)}\cdot\int\limits_{y\in Y}\lambda(x_i,y)g(y)dy$ for $i=1,2$. The change in welfare between these two allocations is:

$$W(\tilde{\lambda})-W(\lambda)\equiv\Delta_W=\int_{\underline{y}}^{\overline{y}}\underbrace{\left[\pi^p(\theta_2)S(x_2,y)-\pi^p(\theta_1)S(x_1,y)\right]}_{\hat{Z}^{CI}(y|\theta)}\varepsilon(y)g(y)dy,$$

where $\theta=(\theta_1,\theta_2)$. Note that $\hat{Z}^{CI}(y|\theta)$ is continuous but not necessarily monotone, provided that $S(x,y)$ is continuous in y for all x.

Now, let Figure D.1 be an arbitrary representation of $\hat{Z}^{CI}(y|\theta)$, and consider $\varepsilon(y)$ defined as follows:

$$\int_{y\in[\underline{y},y_1]} \varepsilon(y)g(y)dy = 0, \quad \int_{y\in[y_1,y_2]} \varepsilon(y)g(y)dy = 0, \text{ and } \int_{y\in[y_2,\overline{y}]} \varepsilon(y)g(y)dy = 0$$

and more importantly, $^{\partial\varepsilon(y)}/_{\partial y} > 0$ for $y \in Y \setminus [y_1, y_2]$ and $^{\partial\varepsilon(y)}/_{\partial y} < 0$ for $y \in [y_1, y_2]$.[36] Thus, the change in welfare is:

$$\int_{y\in[\underline{y},y_1]} \hat{Z}^{CI}(y|\theta)\varepsilon(y)g(y)dy + \int_{y\in[y_1,y_2]} \hat{Z}^{CI}(y|\theta)\varepsilon(y)g(y)dy + \int_{y\in[y_2,\overline{y}]} \hat{Z}^{CI}(y|\theta)\varepsilon(y)g(y)dy.$$

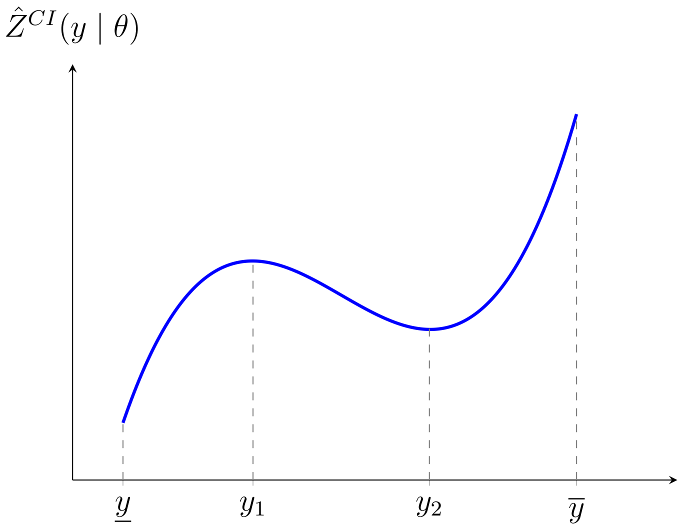

Figure D.1: Change in Welfare - Extension

Since $\hat{Z}^{CI}(y|\theta)$ is monotonically increasing over the interval $[\underline{y}, y_1]$ choosing $\varepsilon(y)$ to be monotonic increasing ensures that the first term above is positive. Similarly, each term can be shown to be positive, which collectively guarantees a welfare improvement over the interior allocation $\lambda(x, y)$.[37] The analysis holds for any interior allocation $\lambda(x, y) \in (0, 1)$ over any arbitrary subset $Y' \subseteq Y$.

Lemma 2 implies the following: "Given an equilibrium market tightness (θ_1^*, θ_2^*), the optimal allocation is always on the corner, that is, $\lambda(x_1, y) \in \{0, 1\}$". Specifically, the following yields the optimal allocation:

[36] Please refer to y_1 and y_2 defined in Figure D.1, here and henceforth.
[37] Notice, one can always find a such an $\varepsilon(y)$ through a very small perturbation around interior $\lambda(x, y)$.

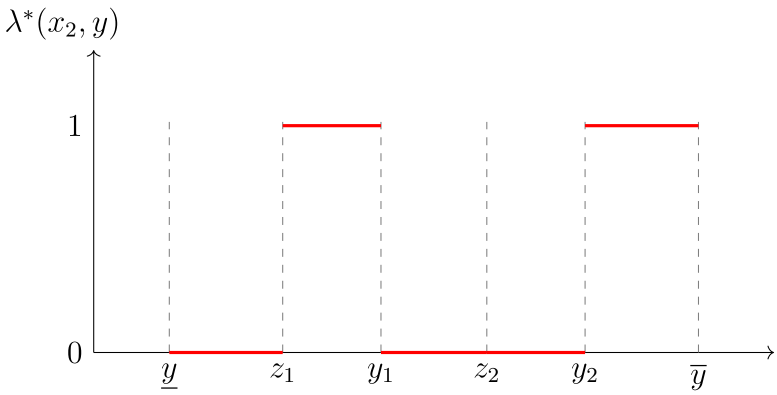

Figure D.2: Optimal allocation rule $\lambda^*(x_2, y)$.

Corollary D.1. *Let $\theta^* \equiv (\theta_1^*, \theta_2^*)$ be the equilibrium market tightness. Suppose $\hat{Z}^{CI}(y|\theta^*)$ is as in Figure D.1. Then the optimal allocation $\lambda^*(x, y)$ is as follows: $\lambda^*(x_1, y) = 1 - \lambda^*(x_2, y)$ and:*

$$\lambda^*(x_2, y) = \begin{cases} 0 & y \in [\underline{y}, z_1] \cup [z_2, z_3] \\ 1 & y \in [z_1, z_2] \cup [z_3, \overline{y}] \end{cases}$$

for some z_1, z_2, z_3 such that $\underline{y} < z_1 < y_1 < z_2 < y_2 < z_3 < \overline{y}$ as in Figure D.2.

Proof. By Lemma 2, a monotone increasing perturbation of $\lambda(y)$ over the interval $[\underline{y}, y_1]$ such that $\int_{y \in [\underline{y}, y_1]} \varepsilon(y) g(y) dy = 0$ guarantees that $\int_{y \in [\underline{y}, y_1]} \hat{Z}^{CI}(y|\theta)\varepsilon(y) g(y) dy > 0$. Since $\varepsilon(y)$ monotone increases and sums up to 0, there exists $z \in (\underline{y}, y_1)$ such that $\varepsilon(y) \leq 0$ if and only if $y \leq z$. This implies $\lambda(x_1, y) \leq \tilde{\lambda}(x_1, y)$ if and only if $y \leq z$. Thus, moving towards PAM only in the interval $[\underline{y}, y_1]$ increases the welfare. Therefore, one can keep increasing the welfare only over the region $[\underline{y}, y_1]$ by trembling as much as possible, which proves that there exists $z_1 \in (\underline{y}, y_1)$ such that $\lambda^*(x_1, y) = 1$ for $y \in [\underline{y}, z_1]$, and $\lambda^*(x_1, y) = 0$ for $y \in [z_1, y_1]$. Analogously, the optimal allocation for other regions follows. $\square$

D.2 Proof of Proposition 5

Notice that, if $\pi^p(\theta^*)S(x, y)$ is super-modular given equilibrium $\theta^* = (\theta_1^*, \theta_2^*)$, then $\hat{Z}^{CI}(y|\theta^*)$ is increasing everywhere. Thus, Corollary D.1 implies that the optimal allocation is such that $\lambda(x_1, y) = 1$ for $y \leq \hat{y}^{PAM}$ for some $\hat{y}^{PAM} \in (\underline{y}, \overline{y})$, and $\lambda(x_1, y) = 0$ otherwise. Moreover, given θ_1^* is the equilibrium market tightness in submarket x_1, $\hat{y}^{PAM}$ is such that $\theta_1^* = G(\hat{y}^{PAM})/1-f(x_2)$. The second part of the Proposition 5 analogously follows.

D.3 Proof of Corollary 4

Recall $\hat{Z}^{CI}(y|\theta) = \pi^p(\theta_2)S(x_2,y) - \pi^p(\theta_1)S(x_1,y)$. One can easily see that $\hat{Z}^{CI}(y|\theta)$ increases in θ_1. Taking derivative of $\hat{Z}^{CI}(y|\theta)$ with respect to y yields the following:

$$\frac{\partial \hat{Z}^{CI}(y|\theta)}{\partial y} \equiv \hat{Z}^{CI}_y(y|\theta) = \pi^p(\theta_2)S_y(x_2,y) - \pi^p(\theta_1)S_y(x_1,y).$$

Since $S_y(y|\theta) > 0$, it follows that $\hat{Z}^{CI}_y(y|\theta)$ also increases in θ_1, and thus assigns its minimum value at $\theta_1 = 0$ and $\theta_2 = {}^1\!/_{f(x_2)}$. That is, $\hat{Z}^{CI}_y(y|\theta_1 = 0, \theta_2 = {}^1\!/_{f(x_2)}) \leq \hat{Z}^{CI}_y(y|\theta)$ for any θ and any y.

Let $\grave{y} := \arg\min_{y \in Y} S_y(x_1,y) - \pi^p\left({}^1\!/_{1-f(x_1)}\right)S_y(x_2,y)$, that is, $\grave{y}$ is the argument at which $\hat{Z}^{CI}_y(y|\theta_1 = 0, \theta_2 = {}^1\!/_{f(x_2)})$ assigns its minimum value. Now, notice the following: $\pi^p\left({}^1\!/_{f(x_2)}\right)S_y(x_2,\grave{y}) - S_y(x_1,\grave{y}) \geq 0$ implies that $\hat{Z}^{CI}_y(y|\theta) \geq 0$ for any θ and any $y \in [\underline{y}, \overline{y}]$. As a result, $\pi^p(\theta^*)S(x,y)$ is super-modular at equilibrium θ^*, and the optimal sorting exhibits PAM by Proposition 5**??**. Therefore, the planner simply optimizes the welfare by solving the following problem:

$$\max_{\hat{y} \in Y} \pi^p\left(\frac{G(\hat{y})}{f(x_1)}\right) \int_{\underline{y}}^{\hat{y}} S(x_1,y)g(y)dy + \pi^p\left(\frac{1-G(\hat{y})}{1-f(x_1)}\right) \int_{\hat{y}}^{\overline{y}} S(x_2,y)g(y)dy,$$

which finishes the proof of Corollary 4 (i). The proof of (ii) follows analogously.

E Improvement in the Meeting Technology

Recall the sorting condition under complete information for a given technology π^p:

$$Z^{CI}(\theta_1|\pi^p) = \pi^p(\theta_2) \underbrace{\left[S(x_2,y_2) - S(x_2,y_1)\right]}_{\Delta S_2} - \pi^p(\theta_1) \underbrace{\left[S(x_1,y_2) - S(x_1,y_1)\right]}_{\Delta S_1}.$$

If the equilibrium sorting is PAM, then the equilibrium θ_1^* is such that $Z^{CI}(\theta_1^*|\pi^p) > Z^{CI}(\overline{\theta}_1|\pi^p) = 0$ and $\theta_1^* > \overline{\theta}_1$. Note that $\left({}^{\pi^p(\overline{\theta}_1)}\!/_{\pi^p(\overline{\theta}_2)}\right) = \left({}^{\Delta S_2}\!/_{\Delta S_1}\right)$.

E.1 Proof of Proposition 6

Take an arbitrary $\hat{\pi}^p \in \mathbf{\Pi}^p$ such that $\hat{\pi}^p \rhd \pi^p$. Notice $Z^{CI}(\bar{\bar{\theta}}_1|\hat{\pi}^p) = 0$ if and only if $\left(\hat{\pi}^p(\bar{\bar{\theta}}_1)/\hat{\pi}^p(\bar{\bar{\theta}}_2)\right) = \left(\Delta S_2/\Delta S_1\right) = \left(\pi^p(\bar{\theta}_1)/\pi^p(\bar{\theta}_2)\right)$. It is easy to verify that:

$$\frac{\hat{\pi}^p(\theta_1)}{\hat{\pi}^p(\theta_2)} \leq \frac{\pi^p(\theta_1)}{\pi^p(\theta_2)} \text{ if and only if } \theta_1 \leq 1,$$

which holds with equality if $\theta_1 = \theta_2 = 1$. In short, given any market tightness θ_1 and $\theta_2 \equiv \left(1-f(x_1)\theta_1\right)/\left(1-f(x_1)\right)$, the ratio of meeting probabilities in submarkets x_1 and x_2, gets flatter as the technology improves (as can be seen in Figure E.1). Thus, a supermodular $S(x, y)$ implies $\bar{\bar{\theta}}_1 < \bar{\theta}_1 < \theta_1^*$.

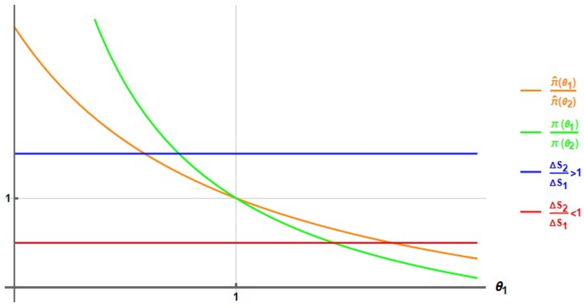

Figure E.1: Monotone Comparative Statics for the Meeting Technology

Recall that the meeting technology gets flatter *everywhere* at an improved technology, that is, let $\partial\pi^p(\theta)/\partial\theta \leq \partial\hat{\pi}^p(\theta)/\partial\theta$ for any $\theta \in [0, \min\{1/f(x_1), 1/1-f(x_1)\}]$ given $f(x_1)$. We also know that $\partial W\left(\lambda^1(x_1),\lambda^2(x_1)\right)/\partial\lambda^k(x_i)$ is monotonically decreasing for any $i = 1, 2$ and any $k = 1, 2$ (see Lemma B.1 in the online supplement). Given π^p, let $\theta_i^* = \min\{\theta_1^*, \theta_2^*\}$ and thus $\theta_i^* \leq 1 \leq \theta_j^*$. Therefore,

$$\frac{\partial W\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial\lambda^k(x_i)}|_{\theta_i=\theta_i^*} = 0 \geq \frac{\partial W\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial\lambda^k(x_i)}|_{\theta_i=1}$$

for some $k = 1, 2$.[38] The inequality becomes strict unless $\min\{\theta_1^*, \theta_2^*\} = \max\{\theta_1^*, \theta_2^*\}$. Suppose that is the case from now on.

Thus, at equal market tightness where the parents-to-children ratio is equal to 1 in both market, the designer would like to allocate some of type-k parents into sub-

[38] For a supermodular $S(x, y)$, consider the corner $(\lambda^1(x_1), \lambda^2(x_1))$ that exhibits PAM and $\theta_1 = 1 = \theta_2$; that is, either $\lambda^1(x_1) = 1$ and $\lambda^2(x_1) \in [0, 1)$, or $\lambda^1(x_1) \in (0, 1)$ and $\lambda^2(x_1) = 0$. Consider analogous λ's for submodular $S(x_i, y_j)$.

market x_j because ${}^{\partial W(\lambda)}/_{\partial \lambda^k(x_i)}|_{\theta_i=1} < 0$. Doing so will have two effects: congestion and decongestion effects, which link to the probability of meeting given the technology, and the surplus effect. Notice the surplus effect is linear, whereas the meeting technology is convex. Hence, decreasing $\lambda^k(x_i)$ at $\theta_i = 1$ increases the probability of meeting in submarket-x_i and decreases in submarket-x_j at different and non-constant rates.

Now, for an improved technology $\hat{\pi}^p$ defined above, the congestion and decongestion effects become less pronounced, leading to less divergence from the equal market tightness $\theta_i = 1 = \theta_j$. Let θ_i^{**} be the equilibrium market tightness with $\hat{\pi}^p$. Therefore, $|1 - \theta_i^{**}| < |1 - \theta_i^*|$, which simply implies $\overline{\overline{\theta_1}} < \theta_1^{**}$. The proof follows analogously for a submodular $S(x, y)$ and NAM.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) for language refinement. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.