Designing the Menu of Licenses for Foster Care *

Altınok, Ahmet † MacDonald, Diana‡

October 2022

Preliminary and Incomplete
Please do not Circulate

Abstract

This paper studies the menu of licenses designed by the child welfare agency to screen foster parents. We develop a two-sided matching model with heterogeneous agents, search frictions, private information, and a designer who coordinates match formation through a menu of contracts. We focus on incentive-compatible contracts, examine optimal transfers, and analyze sorting patterns that arise in equilibrium as in Becker (1973). We establish three main results: **(i)** foster parents in different licenses will never care for the same group of foster children, **(ii)** complementarities do not ensure that Positive Assortative Matching (PAM) will arise in equilibrium, and **(iii)** the equilibrium transfer scheme is not unique and does not affect the *unique* equilibrium sorting pattern. Moreover, our results suggest that an optimal menu of licenses must not only account for the child's attribute (as it is in practice), but also for other characteristics of the market such as parents' types, supply of parents, and supply of children.

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

Keywords : Adverse Selection · Matching · Sorting · Search · Foster Care

---

*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.

†Department of Economics, Nazarbayev University, ahmet.altinok@nu.edu.kz

‡Research Department, Banco de México, diana.macdonald@banxico.org.mx

1 Introduction

Each year more than a half-million children spend at least one day in the US foster care system, a federal program that costs taxpayers almost US$28 billion annually. The foster care system provides out-of-home care for children removed from their homes due to abuse, maltreatment, neglect, or other reasons.[1] While in foster care, children are placed with *foster parents* or in *institutional care*.[2] The former are private individuals licensed to provide 24-hour care for children in a family-based environment, and the latter are licensed facilities that provide 24-hour care for several children at once. For child welfare agencies, placement decisions of children are a key aspect of the market with two main goals: match children with the right foster parent and avoid placing children in institutional care.[3]

Foster care can be viewed as a two-sided matching market, where foster parents have preferences over children, and child welfare agencies have preferences over foster parents (on behalf of children). As in many other markets, matches form in the presence of *private information* and *search frictions*. In words, a foster parent's ability to provide care for a child is unknown to the child welfare agency, and children and parents meet stochastically. Aiming at solving the *adverse selection problem* in the presence of search frictions, the child welfare agency offers a menu of licenses to foster parents who are expected to sort themselves across licenses and reveal their types truthfully. In practice, a license specifies the type of child a parent can foster and the corresponding transfer received by parents. Furthermore, as a rule of thumb, children are grouped by the level of care needed, 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

[1] A child can enter foster care for several reasons such as sexual or physical abuse, parents' drug or alcohol addictions, parents' incarceration, parents' inability to provide care, parents' death, inadequate housing, abandonment, child's behavioral problem, child's drug addiction, or child's alcohol addiction.

[2] Foster parents provide the highest source of out-of-home care. At the end of the federal fiscal year of 2020, the number of children in foster care was 407,493, out of which 83% were placed with foster parents, and 10% were placed in institutional care (U.S. Department of Health and Human Services, 2020).

[3] Regarding the first goal, research shows that children in foster care can be re-traumatized when experiencing a negative placement, resulting in a lower physical and behavioral well-being (Villodas et al., 2016). Related to the second goal, evidence suggests that children placed in institutional care have lower academic outcomes, lower levels of education, a higher risk to engage in delinquent behavior, and a higher risk of criminal convictions during adulthood. (Berrick et al., 1993; Mech et al., 1994; Ryan et al., 2008; Dregan and Gulliford, 2012).

standard needs, while in the latter foster parents can foster children with standard needs and also children with special needs. Parents receive US$20.80 per day for children with standard needs, and US$36.87 for children with special needs. Nowadays, these transfers are based only on the cost of providing care for a child and do not depend on any other characteristics in the market, raising the question of whether the current menu of licenses can achieve its screening objective.

This paper studies the menu of licenses designed by the child welfare agency to screen foster parents. To achieve this objective, we develop a two-sided matching model with heterogeneous agents, search frictions, private information, and a designer who coordinates match formation through a menu of contracts. We focus on incentive-compatible contracts, examine optimal transfers, and analyze sorting patterns that might arise in equilibrium as in Becker (1973).[4] In particular, we establish *three* main results: **(i)** foster parents in different licenses will never care for the same group of foster children, **(ii)** complementarities do not ensure that Positive Assortative Matching (PAM) will arise in equilibrium, and **(iii)** the equilibrium transfer scheme is not unique and does not affect the *unique* equilibrium sorting pattern.

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 their disability status, with (x_1) or without (x_2) a disability; and parents are heterogeneous in their ability to provide care, low (y_1) or high (y_2) ability. A child's attribute is common knowledge, and a parent's attribute is private information. We assume that children receive higher 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 all matches are profitable. As in practice, we divide the market into submarkets for each child's attribute. In other words, there is a submarket populated by children with a disability and another submarket populated by children without a disability. At the beginning of the game, 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 to each submarket, and **(2)** a corresponding transfer when a match forms. After observing the menu of licenses, each parent chooses a license. Next, the randomization device is realized and parents are allocated across submarkets determining endogenously

[4] Becker (1973) analyzes a frictionless two-sided matching market in a marriage context, and provide sufficient conditions in the match payoff function such that the equilibrium sorting exhibits positive or negative assortative matching (PAM or NAM).

the market tightness (parents-to-children ratio) of each submarket. Lastly, meetings take place stochastically, matches are formed, and transfers take place. We assume that, the probability of a child (parent) meeting a parent (child) is represented by a technology function that is increasing (decreasing) and concave (convex) in the market tightness.

We start by analyzing the complete information case where the designer maximizes the expected utility of children minus transfers subject to the interim participation constraints [PC] of parents. After replacing the [PC] in the objective function, the designer's problem reduces to maximize the total expected net utility of matches, where the net utility of a match refers to the child's utility minus the parent's cost, i.e., welfare of a match. We establish results for a super- and a sub-modular net utility function. Now, we discuss the case of super-modularity and leave sub-modularity to be discussed in the body of the paper.

First, if the net utility of a match is super-modular then it is never optimal for the designer to allocate both type-y parents with strictly positive probability into submarkets x_1 and x_2.[5] This result rationalizes the nested nature of the licenses described above for the state of Arizona. Now, to understand the idea behind it, suppose the designer allocates both type of parents into both submarkets with strictly positive probabilities. In a frictionless environment, the designer can increase the total welfare by increasing the probability with which **(i)** y_2-parents are allocated into submarket x_2, and **(ii)** y_1-parents are allocated into submarket x_1. This reallocation across submarkets increases total welfare due to super-modularity because pairs $\{(x_2, y_2), (x_1, y_1)\}$ (PAM) yield a higher net utility than the pairs $\{(x_1, y_2), (x_2, y_1)\}$ (NAM). Hence, the designer has incentives to reassign probabilities until at least one of these hits a corner. By incorporating search frictions, the reallocation mentioned might not increase total welfare because even when the designer modifies probabilities (i) and (ii) by the same amount, the meetings probabilities in each submarket do not necessarily change in the same magnitude. Yet, the designer can still reassign probabilities keeping the same market tightness across submarkets.

Second, we find that super-modularity in the net utility function is neither sufficient nor necessary for the optimal sorting to exhibit PAM. [6] For a frictionless envi-

[5]In other words, 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.

[6]In our framework, the randomization device establishes who can match with whom in the market so we use it to define sorting patterns: a sorting exhibits PAM (NAM) if y_2-parents are allocated to submarket x_2 with a greater (smaller) probability than y_1-parents are. One can equivalently define

ronment with a super-modular net utility function, it is well known that matching agents in a positive assortative way maximizes total welfare. But, when search frictions are introduced, we find that this result does not hold because the expected total welfare, calculated using the meeting technologies in each submarket, is not necessarily super-modular even if the net utility is super-modular. By imposing a lower bound on the fraction of type-x_2 children along with super-modularity, we can ensure that PAM arises in equilibrium. Intuitively, type-y_2 parents are more desirable in any submarket, thus the designer would like to allocate them into the more profitable and thicker submarket x_2. In words, if the faction of children without a disability is high enough relative to the fraction of high ability parents then the designers will always allocate

Third, we find that any transfer scheme that is on the participation constraint for each type of parent is optimal, and it does not affect the equilibrium sorting. Therefore, our framework predicts the same equilibrium sorting regardless of the ex-post or interim participation constraints (for each type of parent). It is intuitive as the parents only care about the expected transfer given a license, which levels the expected cost in equilibrium.

Next, we relax the assumption over the observability of parents' attributes and study the optimal sorting and transfers under the setting with private information. All of the results carry on even if there is private information, although equilibrium sorting may reverse for particular parameter values. Due to the higher expected cost of low-ability parents, the expected transfer they receive is higher than what high-ability parents receive, under the first best menu of licenses. As a result, high-ability parents have incentives to mimic the low-ability ones, and the planner pays information rent to them to eliminate such incentives.

To determine the optimal sorting under private information, one needs to know the cost of a parent-child pairing as well as the parent distribution, which need not be known under complete information.[7] A super-modular (sub-modular) cost function increases the forces for PAM (NAM) at the optimum under information frictions.

---

the sorting pattern through a matching correspondence as standard in the literature, and say that a sorting exhibits PAM if the matching correspondence is a lattice as in (Shimer and Smith, 2000). Since the randomization device provides more information than the correspondence, our sorting 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).

[7] Knowing the net utility is sufficient to determine the equilibrium licences under complete information, we do not need to disentangle utility and the cost to determine the optimal sorting. This is not the case in the presence of information friction.

The intuition is as follows: a super-modular cost function means that the difference between the cost of taking care of a child without a disability for low-ability parents and high-ability parents is higher than the cost of a child with a disability. As a result, the designer would pay less information rent if low-ability parents are allocated to the pool of children with disability. This increases her incentives to sort the market in a positive way. An analogous intuition follows for sub-modular cost function.

Literature Review. The main contribution of this paper is to develop a theoretical matching model with search frictions and adverse selection to study the optimal menu of licenses in the US foster care system. There are a few papers analyzing foster care as a matching market. Slaugh et al. (2015) studies the Pennsylvania Adoption Exchange program, a computational tool created to facilitate the adoption of children in foster care and make several recommendations to improve the success of adoptions. Olberg et al. (2021) constructs a dynamic search and matching model to compare two different search processes use by the child welfare agencies to identify potential adoption matches between parents and children. MacDonald (2022) conducts an empirical analysis that yields four new facts related to match transitions of children in foster care relinquished for adoption. Later on, she develops a dynamic search and matching model where parents and children can form reversible (foster) or irreversible (adoption) matches to rationalize these empirical facts. Lastly, Robinson-Cortés (2019) presents an empirical framework to study how children are assigned to foster homes using a confidential dataset, and uses the estimates to study different policy interventions.

This paper is related to the literature on assortative matching under asymmetric information. In a principal-agent setting with adverse selection, several papers have studied sorting patterns arising from microfinance loan contracts where a population of heterogeneous borrowers optimally match into pairs (Ghatak, 1999; Van Tassel, 1999; Ghatak, 2000; Guttman, 2008; Altinok, 2022). As in our framework, the lender can induce PAM or NAM, but a significant difference is that forming matches has a risk-sharing component. In a principal-agent setting with moral hazard, Serfes (2005) analyzes equilibrium sorting patterns between heterogeneous principals and agents restricting attention to lineal contracts and a CARA utility function. His results rationalize the empirical finding of Ackerberg and Botticini (2002) who document a positive relationship between the degree of risk aversion of tenants and landlords with the riskiness of a crop. Franco et al. (2011) and Kaya and Vereshchagina (2014) examine a framework where a manager assigns heterogeneous workers to teams in the presence of a moral hazard, and show that even in the presence of complementary the equilibrium sorting might exhibit NAM. In a decentralized dynamic setting, Chade (2006) investigates a two-sided matching problem with search and informational frictions where agents make a matching decision based on a noisy signal containing information about the unobservable attribute of the potential partner.

Organization of the Paper. The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents the analysis for the complete information case, and Section 4 extends the analysis to the case of private information. Lastly, Section 5 concludes, and discusses potential extensions. All omitted proofs and examples are in the Appendix.

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 child with a disability, x_2 denotes a child without a disability, and $x_2 > x_1$.The fraction of children with a disability is $f(x_1) \in [0,1]$, whereas the fraction without a disability is $f(x_2) = 1 - f(x_1)$. For the purpose of exposition, we refer the set of children with attribute x as *submarket* x. The other side of the market is constituted by a continuum of *parents* who are heterogeneous in their ability to provide care for a child. In particular, y_1 denotes parents with low ability, y_2 denotes parents with high ability, and $y_2 > y_1$. The fraction of parents with low ability is $g_1 \in [0,1]$, and that with high ability is $g_2 \equiv 1 - g_1$. A parent's ability to provide care is *private information*.

Matches are formed between children and parents, and one-to-one. 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 the parent forms a match with child x. 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 any $x \in X$.

Figure 1 presents two examples of licenses. In Panel 1a, parents holding license $\mathcal{L}$ are allocated to submarket x_1 with probability one, and to submarket x_2 with probability zero. In Panel 1b, parents holding license $\mathcal{L}'$ are allocated to submarket x_1 with probability $\frac{1}{4}$, and to submarket x_2 with probability $\frac{3}{4}$.

All agents are risk-neutral. The designer maximizes children 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 $c(x, y)$.

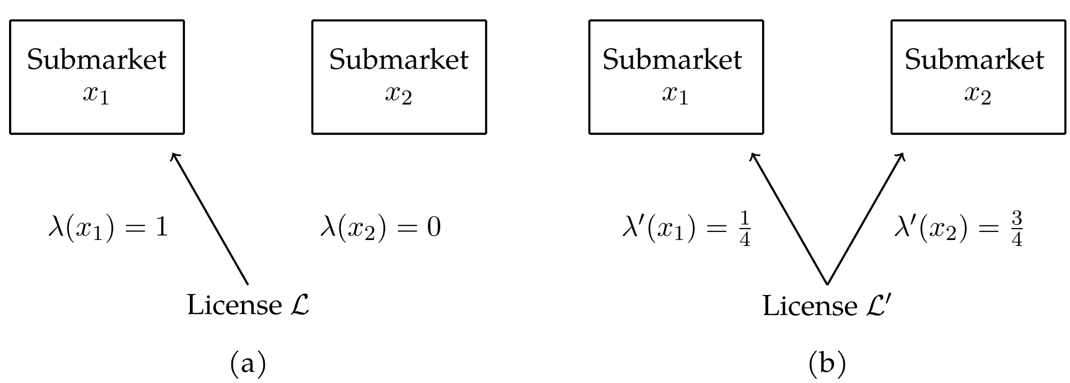
Assumption 1. *(a) for all (x, y), $u(x, y) \geq 0$, $c(x, y) \geq 0$ and $u(x, y) - c(x, y) \geq 0$, (b) $u(x, y)$ is non-decreasing in x and y, and (c) $c(x, y)$ is non-increasing in x and y.*

Assumption 1(a) reflects the following: children are better-off placed with a foster parent than in institutional care, parents always incur a cost when providing care for a child, and all matches are profitable. Assumption 1(b) and 1(c) imply that everyone, absent the transfers, prefers matching with a *high-type* to matching with a *low-type*.

Timing is as follows:

1. First, the designer announces and commits to a menu of licenses. By the revelation principle, we restrict attention to direct revelation mechanisms. Thus, without loss of generality, we consider menus with two licenses, one for each type of parent $\{\mathcal{L}^k\}_{k=1}^2 \equiv \left\{\left\{\left(\lambda^k(x_i), \tau^k(x_i)\right)\right\}_{i=1,2}\right\}_{k=1}^2$.
2. After observing the menu, each parent chooses a license, where $\sigma^y \in \{\mathcal{L}^1, \mathcal{L}^2\}$ denotes this decision. Then, the allocation of parents $\left\{\{\lambda^k(x_i)\}_{i=1,2}\right\}_{k=1}^2$ across submarkets is realized.

Figure 1: Examples of Licenses



3. Next, within each submarket, children and parents 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:

$$\theta_x = \frac{\sum_{k=1}^{2} h^k(y_1)\lambda^k(x) + h^k(y_2)\lambda^k(x)}{f(x)}$$

   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 *meeting technology* $\pi^c(\theta_x)$ where $\pi^c : \mathbb{R}_+ \to [0,1]$ is a strictly increasing and strictly concave function such that $\pi^c(0) = 0$. 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 $\pi^p(\theta_x) = \frac{\pi^c(\theta_x)}{\theta_x}$ and $\pi^p(0) = 1$.

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

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

$$W(x) = \frac{\sum\limits_{k=1}^{2} \Big[\sum\limits_{i=1}^{2} \Big(\big[u(x, y_i) - \tau^k(x) \big] \lambda^k(x) \cdot h^k(y_i) \Big]}{\sum\limits_{k=1}^{2} \lambda^k(x) \cdot \sum\limits_{i=1}^{2} h^k(y_i)}.$$

Then, the designer's problem is:

$$\max_{\left\{ \left\{ \left(\lambda^k(x), \tau^k(x) \right) \right\}_{x \in X} \right\}_{k=1,2}} \left\{ \sum_{x \in X} \pi^c\big(\theta_x\big) \, W(x) \, f(x) \right\} \tag{1}$$

subject to:

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

$$[\text{PC}] \quad \sum_{x\in X} \left[\tau^k(x) - c(x,y_k)\right]\lambda^k(x)\pi^p(\theta_x) \geq 0 \text{ for all } k = 1,2$$

$$[\text{MT}] \quad \theta_x = \frac{1}{f(x)} \sum_{k=1}^{2} \left[\lambda^k(x) \sum_{i=1}^{2} h^k(y_i)\right] \quad \text{for all } x \in X$$

$$[\text{IC}] \quad \sum_{x} \left[\tau^k(x) - c(x,y_k)\right]\lambda^k(x)\pi^p(\theta_x) \geq \sum_{x} \left[\tau^{k'}(x) - c(x,y_k)\right]\lambda^{k'}(x)\pi^p(\theta_x) \text{ for all } k,k'$$

where [FC] is the feasibility constraint and specifies restrictions over each $\lambda^k(x)$ and $\tau^k(x)$. [PC] is the participation constraints guarantees that each parent y_k receives a higher expected payoff when holding license k than when unmatched. The restriction [MT] corresponds to the market tightness (parents-to-children ratio) at each submarket. Lastly, [IC] are the incentive compatibility constraints that ensures that our equilibria are truth-telling.

2.1 Definition of Sorting Patterns

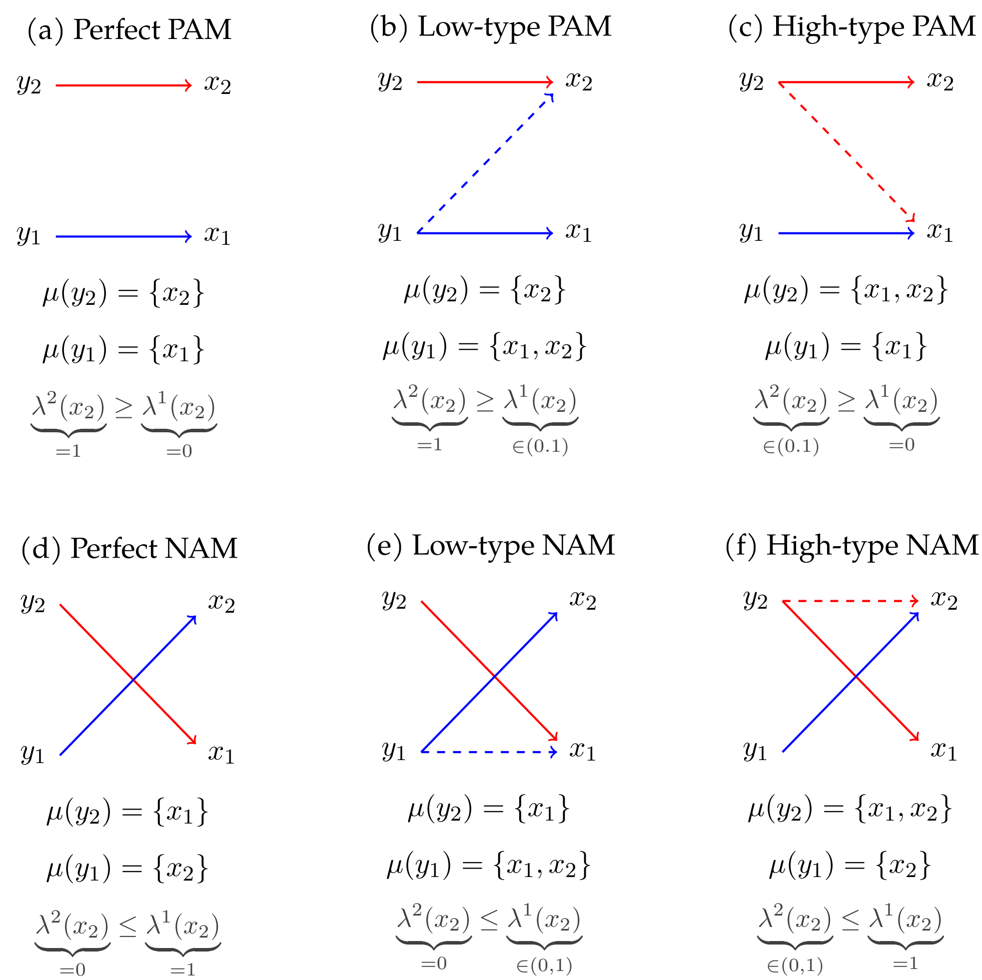
Next, we define a matching correspondence using the randomization device of each license, $\{\lambda^1(x), \lambda^2(x)\}_{x\in X}$, and establish sorting patterns based on the randomization device.

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

We are interested not only in establishing properties that ensures monotone sorting, but also in characterizing the optimal menu of licenses offered by the designer. As a result, our notion of imperfect monotone sorting follows: We say **i–type sorting** if type-y_i parents are allocated to both submarkets x_1 and x_2, while all type-y_{-i} parents are allocated only to submarket $x_j \in \{x_1, x_2\}$.[8] Moreover, we say **i–type PAM** if $-i = j$, and **i–type NAM**, otherwise. To avoid confusion for the rest of the paper, we call it *low*–type if $i = 1$ and *high*–type if $i = 2$. Figure 2 illustrates our notion of monotone sorting patterns.

[8] Here, $-i$ denotes parents of type that is not i. Formally, we say i–type sorting if $\lambda^i(x) \in (0,1)$ while $\lambda^{-i}(x) \in \{0,1\}$.

Figure 2: Perfect & Imperfect Monotone Sorting



3 Equilibrium Analysis: Complete Information

In this section, we present the optimal sorting as well as the equilibrium transfers under complete information. First, note that by incorporating the relevant constraints into the objective function in Equation 1, reduces the designer's problem to the following:

$$\max_{\{\lambda^k(x_1),\lambda^k(x_2)\}_{k=1,2}} \left\{ \sum_{x\in X} \pi^p(\theta_x) \left[\sum_k \Big(\underbrace{u(x, y_k) - c(x, y_k)}_{S(x,y)} \Big) \lambda^k(x) g_k \right] \right\} \tag{2}$$

subject to $\lambda^k(x) \geq 0$ for all (k, x), and $\sum_{x \in X} \lambda^k(x) = 1$ for all $k = 1, 2$. For notational ease, let λ_x^k denote $\lambda^k(x)$, and $U_x^y \equiv U(x, y) = u(x, y) - c(x, y)$ denote the surplus of a match (x, y). Additionally, let x and y take values 1 or 2.

Lemma 1. *For at least one of the licenses, the optimal randomization rule yields a corner solution whenever U_x^y is super- or sub-modular.*

Proof. See Appendix. □

Lemma 1 states that if the net utility of a match is super-or sub-modular then it is never optimal for the designer to allocate both type-y parents with strictly positive probability into submarkets x_1 and x_2. The intuition is as follows. Suppose that the designer allocates both type of parents into both submarkets with strictly positive probabilities. If the net utility of a match is super-modular, then the pairs $\{(x_2, y_2), (x_1, y_1)\}$ (PAM) yield a higher net utility than the pairs $\{(x_1, y_2), (x_2, y_1)\}$ (NAM) do. In a frictionless environment, the designer can increase (by the same amount) the probability with which **(i)** y_2-parents are allocated into submarket x_2, and **(ii)** y_1-parents are allocated into submarket x_1. Such reallocation of parents across submarkets increases the welfare due to supermodularity. Now, under the presence of search frictions, the designer can still reallocate parents across submarkets keeping the market tightness constant. In particular, the designer can do (i) and (ii) but not necessarily by the same amount. Note that, this may not increase welfare because the meetings are stochastic. Yet, if it does, then the designer increases the welfare by doing (i) and (ii) until she cannot do anymore, i.e., at least one probability hits to a corner. If it decreases the welfare, then she trivially increases the welfare by doing a reverse swap, i.e., opposite of (i) and (ii). Either way, she has incentives to segregate the markets as much as possible. It is quite intuitive, because, if some pairs are "more effective" then the other type of pairs, the designer would like to form the former as much as possible, given the market tightness.

To characterize the optimal randomization rule, we follow a nonstandard technique due to the presence of the corner solution. 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 (k, x), let $\{\lambda_x^k\}_{(k,x)}$

be an arbitrary-feasible interior probability generating a total welfare equal to:

$$W(\lambda_1^1, \lambda_1^2) = \pi^p(\theta_1) \cdot \Big(g_1 \lambda_1^1 U_1^1 + (1-g_1)\lambda_1^2 U_1^2\Big) \\ + \pi^p(\theta_2) \cdot \Big(g_1\, (1-\lambda_1^1) U_2^1 + (1-g_1)\, (1-\lambda_1^2) U_2^2\Big)$$

where:

$$\theta_1 = \frac{g_1\lambda_1^1 + (1-g_1)\lambda_1^2}{f_1} \quad \text{and} \quad \theta_2 = \frac{g_1(1-\lambda_1^1) + (1-g_1)(1-\lambda_1^2)}{1-f_1} \tag{3}$$

After trembling λ_1^1 by ε_1 and λ_1^2 by ε_2, such that $\varepsilon_2 \equiv -\frac{\varepsilon_1 g_1}{1-g_1}$ to ensure that the market tightness in each market remains constant, the change in welfare is equal to:

$$\begin{aligned} \Delta_W &= W(\lambda_1^1 + \varepsilon_1, \lambda_1^2 + \varepsilon_2) - W(\lambda_1^1 \lambda_1^2) \\ &= \varepsilon_1 g_1 \underbrace{\Big(\pi^p(\theta_2)\, [U_2^2 - U_2^1] - \pi^p(\theta_1)\, [U_1^2 - U_1^1]\Big)}_{Z^{CI}(\theta_1)} \end{aligned}$$

where θ_x is defined as in Equation 3.[9] It is easy to see that the designer can always increase total welfare by changing $\{\lambda_x^k\}_{(k,x)}$ such that the market tightness remains constant. The optimal allocation of parents can be characterized by $Z^{CI}(\theta_1)$, and the sign of $Z^{CI}(\theta_1)$ determines the equilibrium sorting.[10] In particular, let $\bar{\theta}_1$ be such that $Z^{CI}(\bar{\theta}_1) = 0$, then the following characterizes the equilibrium sorting:

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.*

Proof. See Appendix. □

Note that, $X(\theta_1)$ represents the difference in expected cost-net utilities between a PAM-pair and a NAM-pair. Therefore, if the equilibrium market tightness is such that $X(\theta_1^*)$ is positive, it means that the expected cost-net utilities of a PAM-pair is higher than that of a NAM pair, leading to PAM at the optimum. Analogously, lower equilibrium market tightness leads optimal sorting towards NAM. An immediate corollary follows:

[9] Notice that $\theta_2 = \frac{1-f_1\theta_1}{1-f_1}$, thus $X(\cdot)$ can be written as a function of only θ_1.

[10] Unless $X(\theta_1) = 0$. But, we show that the designer can increase the welfare by changing either λ_1^1 or λ_1^2 whenever $X(\theta_1) = 0$.

Corollary 1. *(**i**) If $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2-U_1^1}{U_2^2-U_2^1}$ then the equilibrium exhibits (perfect or partial) PAM. (**ii**) If $\pi^p\left(\frac{1}{f_1}\right) \geq \frac{U_2^2-U_2^1}{U_1^2-U_1^1}$ then the equilibrium exhibits (perfect or partial) NAM.*

Proof. See Appendix. □

For a more detail characterization, see Appendix A.[11] It is important to highlight that supermodularity (submodularity) of U_x^y is not a sufficient condition for PAM (NAM) to arise in equilibrium. Recall that, in an environment without search frictions, these conditions on the payoffs suffices.

Now, we illustrate two examples. Example 3 visualizes the optimal parent allocation across submarkets. Example 3 illustrates two environments where, in equilibrium, neither super-modularity implies PAM nor sub-modularity implies NAM. It also illustrates sufficient conditions for monotone sorting described in Corollary 1.

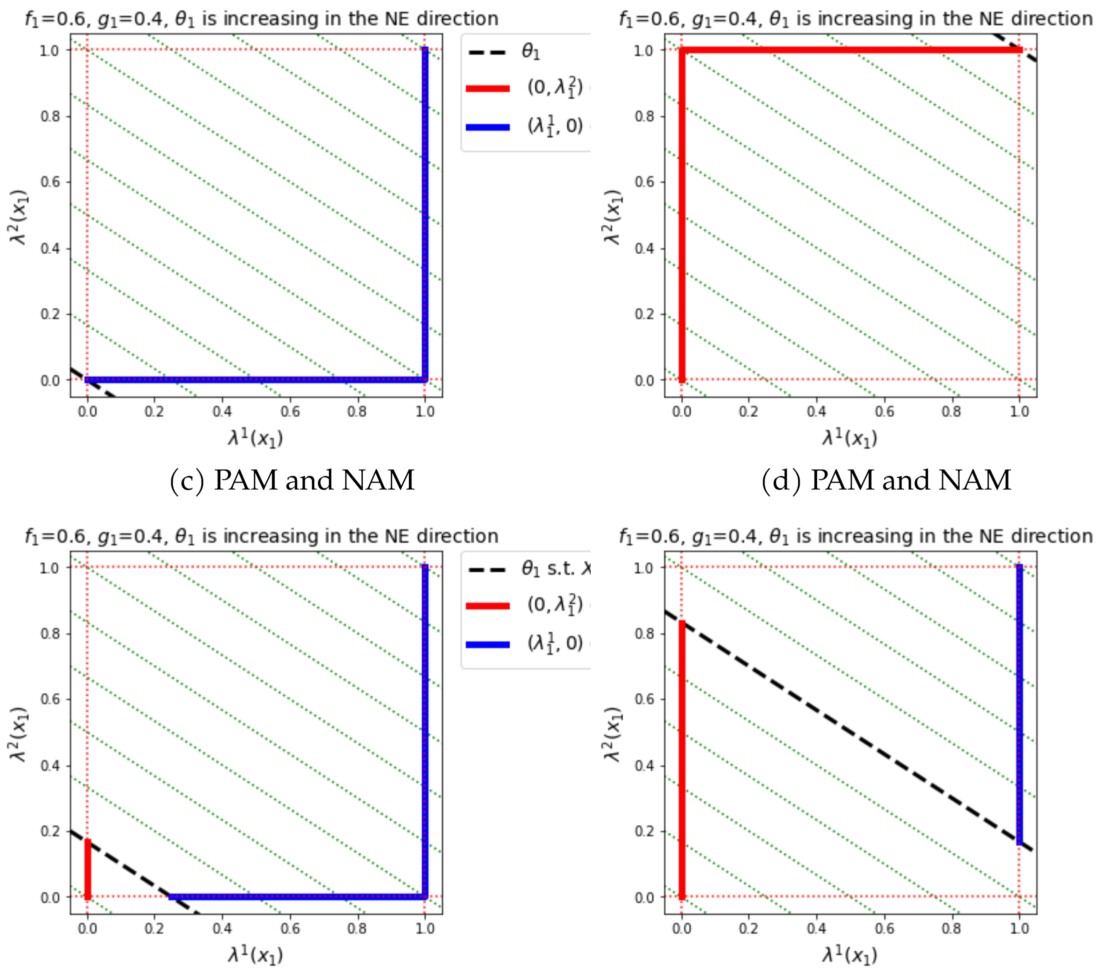
Example 1. Let $f_1 = 0.6$ be the share of children with a disability, and $g_1 = 0.4$ be the share of parents with low ability. Figure 3 presents four different environments with Panels 3a and 3b satisfying the sufficient conditions from Corollary 1, while the environments in Panels 3c and 3d do not. In each graph, 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 $(\lambda_1^1, \lambda_1^2)$ is a feasible allocation of parents. In addition, each black-dotted line corresponds to the values of $(\lambda_1^1, \lambda_1^2)$ such that $X(\overline{\theta}_1) = 0$ where $\overline{\theta}_1 = \frac{\lambda_1^1 g_1 + \lambda_1^2(1-g_1)}{f_1}$. The blue-lines shows the feasible solutions for $X(\theta_1) > 0$ (above the black-dotted line), and the red-lines shows the feasible solutions for $X(\theta_1) < 0$ (below the black-dotted line).

In Panel 3a, the equilibrium sorting only exhibits PAM. The horizontal blue-line corresponds to $(\lambda_1^1, \lambda_1^2)$-points that result in a low-type partial PAM, while the vertical blue-line is a high-type partial PAM. The point $(\lambda_1^1, \lambda_1^2)$ where these two lines intersect corresponds to a perfect PAM. Analogously, in Panel 3b, the horizontal red-line corresponds to $(\lambda_1^1, \lambda_1^2)$-points that result in a low-type partial NAM, the vertical-red lines is a high-type partial NAM, and the intersection is perfect NAM.

Panels 3c and 3d show cases where either PAM or NAM can arise in equilibrium. In the former, there are four possible sorting patterns. In addition to the ones mentioned for Panel, there is given by the vertical red-line. □

[11] Another characterization is given for monotone sorting in terms of the distribution of children. See Corollary 5 in the Appendix.

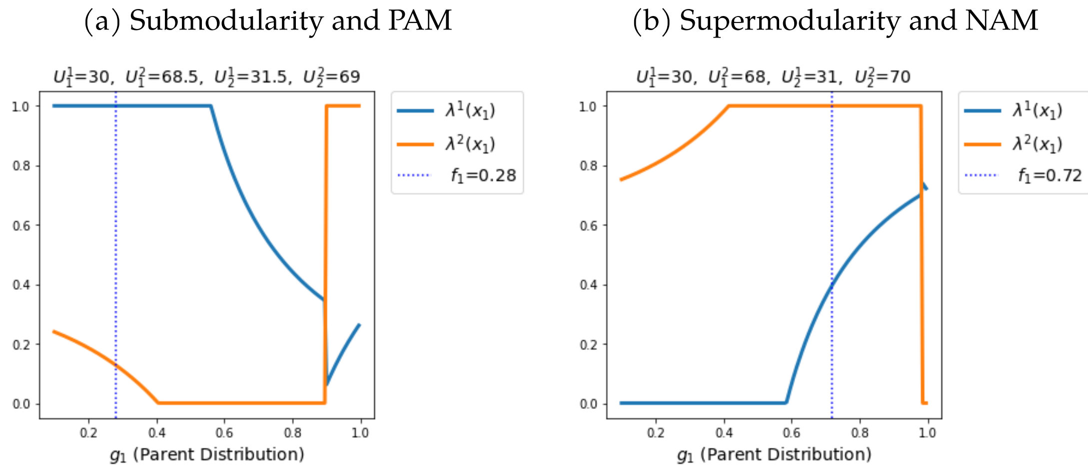
Figure 3: Equilibrium Exhibits Corner Solutions



(a) Only PAM (b) Only NAM (c) PAM and NAM (d) PAM and NAM

Example 2. Figure 4 illustrates environments where sub-modularity (Panel 4a) and super-modularity (Panel 4b) in the cost-net utility function U_x^y are not sufficient for NAM and PAM, respectively. In both panels, we plot the optimal randomization rule for licenses 1 and 2 as a function of each possible value of the share of parents with low ability (g_1). Specifically, we plot the probability with which parents with low (blue line) and high (orange line) ability are allocated into submarket x_1. In Panel 4a, we assume that the share of children with a disability is equal to $0,28$, and U_x^y is a sub-modular function with values $U_2^2 = 69$, $U_1^2 = 68,5$, $U_2^1 = 31,5$ and $U_1^1 = 30$. In Panel 4b, we assume that the share of children with a disability is equal to $0,72$, and U_x^y is a super-modular function with values $U_2^2 = 70$, $U_1^2 = 68$, $U_2^1 = 31$ and

Figure 4: Monotone Sorting Fails



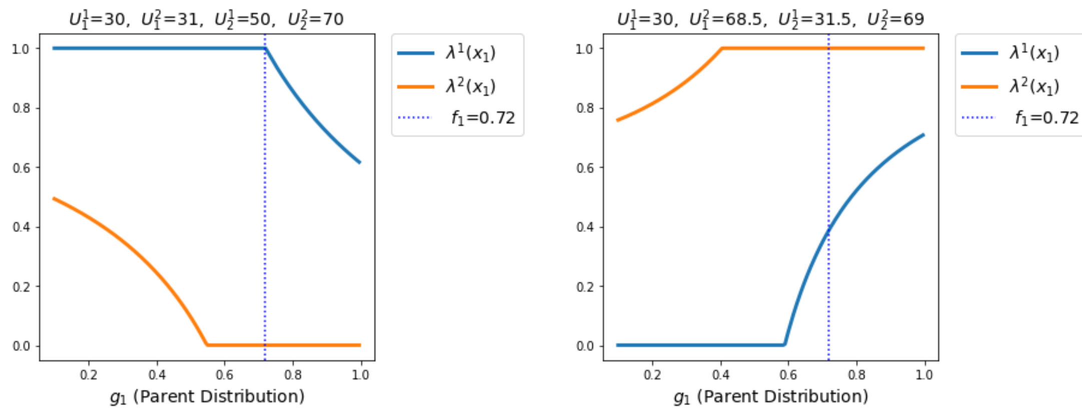
$U_1^1 = 30$. In addition, we assume that the functional form of the meeting technology is $\pi^p(\theta) = \frac{1}{1+\theta}$.

In Panel 4a, we observe the following: **(i)** If $g_1 = 0,2$, then parents with low ability are allocated into submatket x_1 with probability one, and parents with high ability are allocated into both submarkets with a strictly positive probability. This corresponds to a high-type partial PAM. **(ii)** If $g_1 = 0,5$, then low ability parents are allocated only into submarket x_1, and high ability parents are allocated only into submarket x_2. This is perfect PAM. **(iii)** If $g_1 = 0,8$, then parents with high ability are allocated into submarket x_2 with probability one, and parents with low ability are allocated into both submarkets with s strictly positive probability. This corresponds to a low-type partial PAM. **(iv)** As g_1 approaches to one, the sorting pattern reverses and becomes NAM. Similarly, in Panel 4b, the sorting exhibits NAM for most of the possible values of g_1, except for when it approaches to one, in which case the sorting pattern reverses and becomes PAM.

Figure 5 considers environments that satisfy the additional conditions presented in Corollary 1 to ensure monotone sorting. In Panel 5a, we assume that the share of children with a disability is equal to $0,72$, and U_x^y is a super-modular function with values $U_2^2 = 70$, $U_1^2 = 31$, $U_2^1 = 50$ and $U_1^1 = 30$. In this case, the condition over primitives presented in Corollary 1(i) corresponds to $0,22 = \frac{1-f_1}{2-f_1} \geq \frac{U_1^2 - U_1^1}{U_2^2 - U_2^1} = 0,05$. In Panel 5b, we assume that the share of children with a disability is equal to $0,72$, and U_x^y is a sub-modular function with values $U_2^2 = 69$, $U_1^2 = 68,5$, $U_2^1 = 31,5$ and $U_1^1 = 30$. Here, the condition over primitives presented in Corollary 1(ii) corresponds

Figure 5: Monotone Sorting Holds

(a) Supermodularity and PAM (b) Submodularity and NAM



to $0,42 = \frac{f_1}{1+f_1} \geq \frac{U_2^2 - U_2^1}{U_1^2 - U_1^1} = 0,97$.

Lastly, note that there is an interval of g_1 for which the allocation of parents is constant even though g_1 changes. That occurs when we have perfect PAM or NAM. However, if g_1 moves over the region in which equilibrium exhibits partial sorting, then the allocation (either $\lambda^1(x_1)$ or $\lambda^1(x_1)$) also changes. The reason is as follows: If g_1 moves over the partial sorting region, the cost and benefit of allocating a type of parent also changes and as a response equilibrium interior $\lambda^y(x_1) \in (0,1)$ changes. However, the cost and benefit analyses switches from parent y to parent $-y$ when interior $\lambda^y(x_1)$ hits to a corner, which shifts everything. Therefore, a larger change in g_1 is necessary to move from perfect sorting to an interior point–partial sorting. □

Next, we study the optimal transfer scheme. By fixing the optimal allocations $\{\lambda^1(x)\}_{x\in X}$ and $\{\lambda^2(x)\}_{x\in X}$ from Equation 2, the designer solves the following:

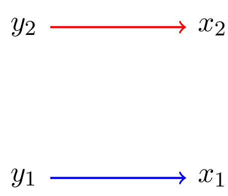
$$\min_{\{\tau^k(x_1),\tau^k(x_2)\}_{k=1,2}} \left\{ \sum_x \pi^p(\theta_x) \sum_k \tau^k(x)\lambda^k(x)g_k \right\} \tag{4}$$

subject to [FC] and [PC] from Equation 1. Recall that the optimal allocation of at least one type of parent is a corner solution, in which case the transfer can be trivially pint-down. Fixing k, if $\lambda^k(x_1) = 1$ then $\tau^k(x_1) = c(x_1, y_k)$, and if $\lambda^k(x_1) = 0$ then $\tau^k(x_2) = c(x_2, y_k)$. Now, in case of an interior solution $\lambda^k(x_1) \in (0,1)$, the transfers $\tau^k(x_1)$ and $\tau^k(x_2)$ satisfy the [PC]. Formally:

Proposition 2. *Given an equilibrium allocation of parents $\{\lambda_x^k\}_{(x,y)}$, any feasible transfer*

Figure 6: Optimal Transfers for PAM

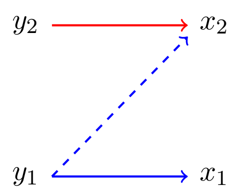
(a) Perfect PAM



$\tau^2(x_2) = c(x_2, y_2)$

$\tau^1(x_1) = c(x_1, y_1)$

(b) Low-type Partial PAM

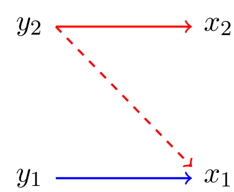


$$\tau^2(x_2) = c(x_2, y_2)$$

$$\tau^1(x_2) \geq 0$$

$$\tau^1(x_1) = \frac{\sum_{i=1}^{2} \pi^p(\theta_i)c(x_i, y_1)\lambda_i^1 - \pi^p(\theta_2)\tau^1(x_2)\lambda_2^1}{\pi^p(\theta_1)\lambda_1^1}$$

(c) High-type Partial PAM



$$\tau^2(x_2) = \frac{\sum_{i=1}^{2} \pi^p(\theta_i)c(x_i, y_2)\lambda_i^2 - \pi^p(\theta_1)\tau^2(x_1)\lambda_1^2}{\pi^p(\theta_2)\lambda_2^2}$$

$$\tau^2(x_1) \geq 0$$

$$\tau^1(x_1) = c(x_1, y_1)$$

schedule for which the participation constraints hold with equality is an equilibrium.

Figure 6 exhibits the optimal transfers for (partial and perfect) PAM. As we can see from Panel 6a, when markets are completely segregated, parents are compensated by exactly the cost of providing care. Panels 6b and 6c, present the transfer for low- and high-type Partial PAM, respectively. Here, we can see that type-y parents who provide care only in submarket x receive the cost of providing care $c(x, y)$; while for type-y parents who provide care in both submarkets there is not a unique solution. Now, as in practice, lets suppose that we include a restriction imposing that parents who provide care in the same market receive the same transfer. Hence, in panel 6b, the transfers are $\tau^1(x_2) = \tau^2(x_2) = c(x_2, y_2)$ and $\tau^1(x_1) = c(x_1, y_1) + \big[c(x_2, y_1) - c(x_2, y_2)\big]\frac{\pi^p(\theta_2)\lambda_2^1}{\pi^p(\theta_1)\lambda_1^1}$. In panel 6c, the transfers are $\tau^1(x_1) = \tau^2(x_1) = c(x_1, y_1)$ and $\tau^2(x_2) = c(x_2, y_2) - \big[c(x_1, y_1) - c(x_1, y_2)\big]\frac{\pi^p(\theta_1)\lambda_1^2}{\pi^p(\theta_2)\lambda_2^2}$.

4 Equilibrium Analysis: Private Information

In this section, we analyze the case where a parent's ability is private information. By assumption, high-ability parents incur in a smaller cost when proving care than low-ability parents, regardless of the disability status of the child. This translates to high-ability parents receiving a smaller expected transfer under the menu of licenses specified under complete information. Thus, in the presence of private information, high-ability parents have incentives to mimic low-ability parents in order to receive greater transfers. This incentive holds for any sorting pattern. Hence, the designer pays informational rents to high-ability parents.

We start our analysis by establishing that Lemma 1 presented in the previous section holds under private information.

Lemma 2. *Even with information frictions, for at least one of the licenses, the optimal randomization rule yields a corner solution whenever U_x^y is super- or sub-modular.*

Proof. See Appendix. □

Now, following the same argument as in the previous section, the condition that characterizes the optimal allocation is equal to:

$$Z^{PI}(\theta_1) = \pi^p(\theta_2) \cdot \left(U_2^2 - U_2^1 + \frac{g_2}{g_1}\Delta c_2\right) - \pi^p(\theta_1) \cdot \left(U_1^2 - U_1^1 + \frac{g_2}{g_1}\Delta c_1\right)$$

where $g_2 = 1 - g_1$, and $\Delta c_x = c(x, y_1) - c(x, y_2) > 0$. Let $\hat{\theta}_1$ be such that $Z^{PI}(\hat{\theta}_1) = 0$. The following characterizes the equilibrium sorting under the incomplete information:

Proposition 3. *Let $\theta_1^{\star\star}$ be the equilibrium market tightness.* ***(i)*** *If $\theta_1^{\star\star} > \hat{\theta}_1$ then the equilibrium exhibits PAM.* ***(ii)*** *If $\theta_1^{\star\star} < \hat{\theta}_1$ then the equilibrium exhibits NAM.* ***(iii)*** *$\theta_1^{\star\star} = \hat{\theta}_1$ is never optimal.*

Proof. See Appendix. □

As in the case of complete information, $\hat{X}(\theta_1)$ represents the difference in expected cost-net utilities between a PAM-pair and a NAM-pair under incomplete information. As a result, the equilibrium sorting is characterized by θ_1^{**} relative to $\hat{\theta}_1$. An immediate corollary follows:

Corollary 2. ***(i)*** *If* $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2-U_1^1+\frac{g_2}{g_1}\Delta c_1}{U_2^2-U_2^1+\frac{g_2}{g_1}\Delta c_2}$ *then the equilibrium exhibits (perfect or partial) PAM.* ***(ii)*** *If* $\pi^p\left(\frac{1}{f_1}\right) \geq \frac{U_2^2-U_2^1+\frac{g_2}{g_1}\Delta c_2}{U_1^2-U_1^1+\frac{g_2}{g_1}\Delta c_1}$ *then the equilibrium exhibits (perfect or partial) NAM.*

Proof. See Appendix. □

Not that the visual illustration of corner solutions in Figure 3 carries over to the case of incomplete information. However, the market tightness that satisfies $\hat{X}(\theta_1) = 0$ may differ, which can potentially change the equilibrium sorting.

Recall, $U_x^y = u(x,y) - c(x,y)$. Let u_x^y and c_x^y denote $u(x,y)$ and $c(x,y)$, respectively. Therefore, one can write $\hat{X}(\theta_1)$ also as:

$$\hat{X}(\theta) = \pi^p(\theta_2) \cdot \left(\left(u_2^2 - \frac{c_2^2}{g_1}\right) - \left(\left(u_2^1 - \frac{c_2^1}{g_1}\right)\right) - \pi(\theta_1) \cdot \left(\left(u_1^2 - \frac{c_1^2}{g_1}\right) - \left(\left(u_1^1 - \frac{c_1^1}{g_1}\right)\right).$$

Recall the sorting condition of the complete information case:

$$X(\theta) = \pi^p(\theta_2) \cdot \left(\left(u_2^2 - c_2^2\right) - \left(\left(u_2^1 - c_2^1\right)\right) - \pi^p(\theta_1) \cdot \left(\left(u_1^2 - c_1^2\right) - \left(\left(u_1^1 - c_1^1\right)\right).$$

By adjusting the cost due to the information frictions (which is higher than the actual cost), the sorting conditions of complete and incomplete information become identical. As g_1 increases, the adjusted cost decreases and approaches to the actual cost. Higher g_1 means less proportion of high-type parents, and thus, less information rents to be distributed. Therefore the equilibrium approaches to the first best. The intuition follows for the other way around if g_1 decreases.

Unlike complete information, the cost-net benefit of a match is not a sufficient statistics to characterize the optimal sorting under the presence of information frictions. One needs to know the cost of a match for characterization.

The following introduces further restrictions for the same sorting pattern to carry over to the case of incomplete information:

Corollary 3. ***(i)*** *If* $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2-U_1^1}{U_2^2-U_2^1} \geq \frac{\Delta c_1}{\Delta c_2}$ *then the equilibrium exhibits (perfect or partial) PAM, and* ***(ii)*** *If* $\pi^p\left(\frac{1}{f_1}\right) \geq \frac{U_2^2-U_2^1}{U_1^2-U_1^1} \geq \frac{\Delta c_2}{\Delta c_1}$ *then the equilibrium exhibits (perfect or partial) NAM, regardless of the existence of information frictions.*

This simply follows from Corollaries 1 and 2. It, for example, implies that if cost-net utility is super-modular to a sufficiently strong degree, and the cost function displays even stronger super-modularity, then the optimal sorting exhibits PAM regardless of the observability of parents' characteristics.

The intuition is as follows: Suppose that the cost function $c(x, y)$ is super-modular (to a *greater extent* than the submodularity of cost-net utility). This means, the difference in cost of taking care of a child between high- and low-type parents is higher in the submarket of children without disability. Notice these differences are negative. Therefore, the information rent to be distributed becomes larger if high-type parents mimic the low-type ones to be in the submarket of children with disability (because $|c(x_1, y_2)-c(x_1, y_1)| > |c(x_2, y_2)-c(x_2, y_1)|$). It encourages the designer to allocate low-type parents more into submarket of children without disability, exhibiting NAM. An analogous intuition follows for sub-modularity and NAM (Part (ii) of Corollary 3). For a further characterization of the private information setting, see Appendix B.

Following example illustrates the equilibrium sorting of the environment in Example 3 (corresponding to Figure 5) under presence of information frictions. We take the environment where sufficient conditions described in Corollary 1 for monotone sorting are satisfied. Then, we introduce cost functions in such a way that the same sorting carries over to the case of incomplete information (Figure 7). That is, cost function shows stronger sub-modularity (left chart in Figure 7) and super-modularity (right chart in Figure 7) than the cost-net utility does. This implies that the conditions described in corollary 3 are satisfied, and thus the equilibrium sorting remains unchanged. We also introduce cost functions that violates conditions described in corollary 3 to the same environments, and illustrate regions of a parameter on which sorting reverses under the presence of information frictions.

Example 3. Figure 7 illustrates the equilibrium sorting of the environment in Example 3 (corresponding to Figure 5) under presence of information frictions. The dashed lines are equilibrium sorting under incomplete information (denoted by $\hat{\lambda}^y(x)$) while the solid ones are first best sorting. Given such values of cost function $c(x, y)$ in Figure 7, the equilibrium sorting—qualitatively—remains the same as it is under the complete information setting. However, notice that the value(s) may differ from their counterparts under the complete information.

The story is different though for such values of cost function introduced in Figure 8. When information frictions exists, the equilibrium sorting switches from PAM to NAM (on the right chart), and from NAM to PAM (on the left chart) for lower fractions of low-type parents, as expected. Recall that the "adjusted" cost is $c(x, y)/g_1$ and

Figure 7: Private Information-Sufficiency

(a) (b)

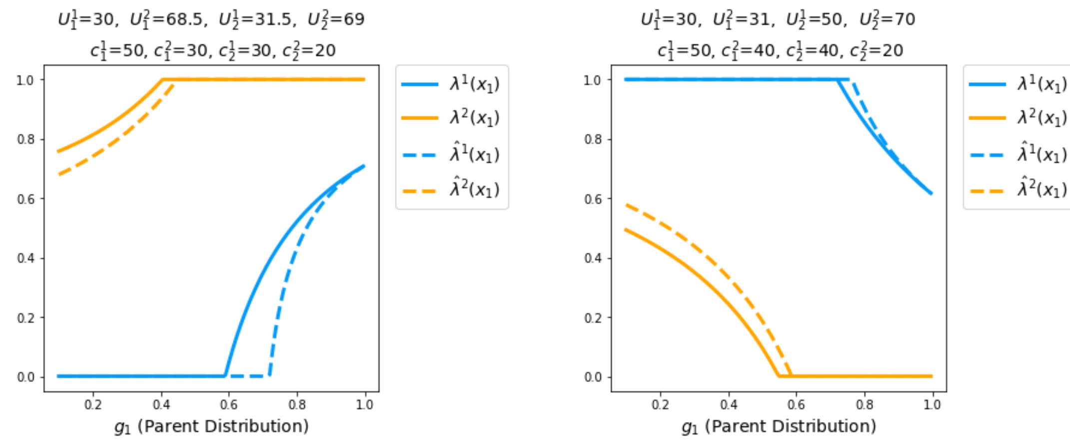
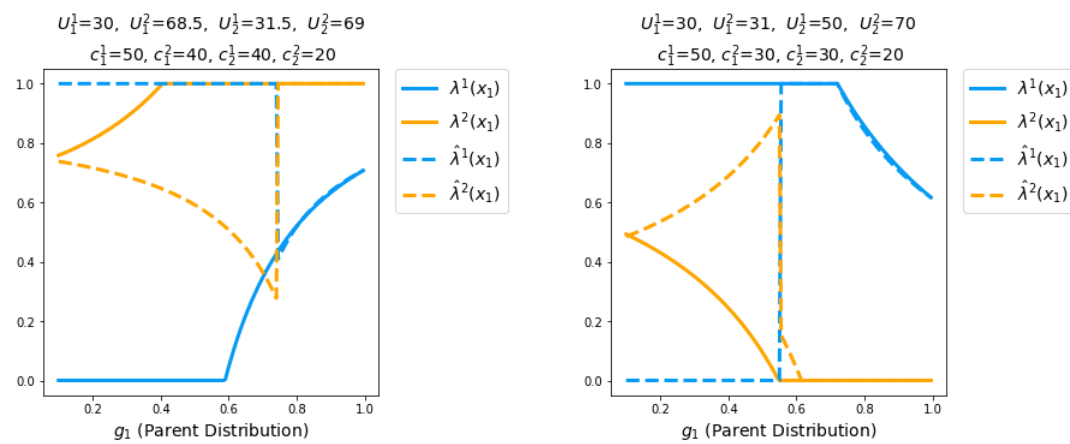


Figure 8: Private Information-Insufficiency

(a) (b)



approaches to the actual one as g_1 increases. Thus, the qualitative change in equilibrium sorting is expected to arise when g_1 is low enough. Otherwise, one would only observe quantitative changes on the equilibrium sorting once information friction is introduced. □

As in the complete information, transfer schedule under private information is also determined by the constraints. Participation constraint of low-type parents as well as the incentive-compatibility constraint of high-type parents determine the equilibrium transfer schedule. Note that, the designers problem, given the optimal allocation of parents into submarkets; $\{\lambda^k(x)\}_{\{k=1,2,\ x\in X\}}$, reduces to the following:

$$\min_{\left\{\left(\tau^k(x)\right)_x\right\}_{k=1}^2} \left\{ \sum_x \pi^p(\theta_x) \sum_k \tau^k(x)\lambda^k(x)g(y_k) \right\} \tag{5}$$

subject to FCs and PCs and ICs, as in complete information case. For an easier notation, let π_i^p, τ_i^k, λ_i^k, and c_i^k denote $\pi^p(\theta_{x_i})$, $\tau^{y_k}(x_i)$, $\lambda^{y_k}(x_i)$, and $c(x_i, y_k)$, respectively. Knowing PC[1] as well as IC[2] holds by equality, these two further simplifies to:

$$\text{PC}[1] : \tau_1^1\lambda_1^1\pi_1^p + \tau_2^1\lambda_2^1\pi_2^p = c_1^1\lambda_1^1\pi_1^p + c_2^1\lambda_2^1\pi_2^p, \quad \text{and}$$

$$\text{IC}[2] : \tau_1^2\lambda_1^2\pi_1^p + \tau_2^2\lambda_2^2\pi_2^p - c_1^2\lambda_1^2\pi_1^p - c_2^2\lambda_2^2\pi_2^p = \underbrace{\tau_1^1\lambda_1^1\pi_1^p + \tau_2^1\lambda_2^1\pi_2^p}_{c_1^1\lambda_1^1\pi_1^p + c_2^1\lambda_2^1\pi_2^p} - c_1^2\lambda_1^1\pi_1^p - c_2^2\lambda_2^1\pi_2^p$$

which then becomes the following

$$\tau_1^2\lambda_1^2\pi_1^p + \tau_2^2\lambda_2^2\pi_2^p = c_1^2\lambda_1^2\pi_1^p + c_2^2\lambda_2^2\pi_2^p + (c_1^1 - c_1^2)\lambda_1^1\pi_1^p + (c_2^1 - c_2^2)\lambda_2^1\pi_2^p.$$

Incorporating these two into designers minimization problem in (5) becomes free of choice variables τ_i^k, which paves out way to the following characterization:

Proposition 4. *Given equilibrium allocation of parents $\{\lambda_x^k\}$, any feasible transfer schedule for which the participation constraint of low-type parent as well as the incentive-compatibility constraint of high-type parent hold by equality is an equilibrium.*

5 Concluding Remarks

This paper analyzes the Foster Care system in the US as a two-sided matching market wherein one side consists of children who are heterogeneous in their disability status, and the other side consists of parents who differ from each other in their ability to take care of a child. We solve for the optimal menu of licenses —which reveals the allocation of parents into subcategories of children as well as the equilibrium transfers that parents receive— under the presence of search frictions. The optimal allocation induces an optimal sorting. Initially, we tackle the problem under the complete information as a benchmark and later extend our analyses to an environment under which parent characteristics are private information.

There are two main results of the paper that hold regardless of the information frictions. (i) It is not optimal to mix multiple types of parents into multiple subcategories of children, and (ii) Super-modularity and sub-modularity of the net utility function are neither sufficient nor necessary for optimal sorting to exhibit positive assortative matching (PAM) and negative assortative matching (NAM), respectively. The former rationalizes the nested licenses in the Foster Care system offered by various states in the US, such as Arizona. The latter has implications on the optimal allocation of parents: Even if the net utility shows complementarity (substitutability) in child and parent characteristics, allocating parents into subcategories such that the sorting exhibits PAM (NAM) is not necessarily optimal due to search frictions.

We also make inferences about the transition of sorting pattern once information friction is introduced: As the share of low-type parents increases, the allocation of parents approaches to the first-best one —the one under the complete information. Because, high-type parents mimic the low-type ones to receive a higher expected transfer. As a result, the designer pays information rent to high-type parents to overcome such incentives. The lower the share of high-type parents, the less the designer cares about such mimicking incentives. However, if the proportion of high-type parents is high, then not only the allocation diverge from the first-best one, but also the optimal sorting may reverse.

References

Ackerberg, D. and M. Botticini (2002): "Endogenous Matching and the Empirical Determinants of Contract Form," *The Journal of Political Economy*, 110, 564–592.

Altinok, A. (2022): "Group Lending, Sorting, and Risk Sharing," *Working Paper*.

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

Berrick, J., M. Courtney, and R. Barth (1993): "Specialized Foster Care and Group Home Care: Similarities and Differences in the Characteristics of Children in Care," *Children and Youth Services Review*, 15, 453–473.

Chade, H. (2006): "Matching with Noise and the Acceptance Curse," *Journal of Economic Theory*, 129, 81–113.

Damiano, E. and H. Li (2005): "Price Discrimination and Efficient Matching," *Economic Theory*, 30, 243–263.

Dregan, A. and M. Gulliford (2012): "Foster Care, Residential Care and Public Care Placement Patterns are Associated with Adult Life Trajectories: Population-based Cohort Study," *Social Psychiatry and Psychiatric Epidemiology*, 47, 1517–1526.

Franco, A., M. Mitchell, and G. Vereshchagina (2011): "Incentives and the Structure of Teams," *The Journal of Economic Theory*, 146, 2307–2332.

Ghatak, M. (1999): "Group Lending, Local Information, and Peer Selection," *Journal of Development Economics*, 60, 27–50.

——— (2000): "Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect," *Economic Journal*, 110, 601–631.

Guerrieri, V., R. Shimer, and R. Wright (2010): "Adverse Selection in Competitive Search Equilibrium," *Econometrica*, 78, 1823–1862.

Guttman, J. (2008): "Assortative Matching, Adverse Selection, and Group Lending," *Journal of Development Economics*, 81, 51–56.

Kaya, A. and G. Vereshchagina (2014): "Partnerships versus Corporations: Moral Hazard, Sorting, and Ownership Structure," *American Economic Review*, 104, 291–307.

Legros, P. and A. Newman (2007): "Beauty is a Beast, Frog is a Prince: Assortative Matching with Non-Tranferabilities," *Econometrica*, 75, 1073–1102.

MacDonald, D. (2022): "Foster Care: A Dynamic Matching Approach," *Working Paper*.

McAfee, P. (2002): "Coarse Matching," *Econometrica*, 70, 2025–2034.

Mech, E., C. Ludy-Dobson, and F. Hulseman (1994): "Life-skills Knowledge: A Survey of Foster Adolescents in Three Placement Settings," *Children and Youth Services Review*, 16, 181–200.

Olberg, N., L. Dierks, S. Seuken, V. W. Slaugh, and U. Unver (2021): "Search and Matching for Adoption from Foster Care," *Working Paper*.

Robinson-Cortés, A. (2019): "Who Gets Placed Where and Why? An Empirical Framework for Foster Care Placement," *Working Paper*.

Ryan, J., J.-M. Marshall, D. Herz, and P. Hernandez (2008): "Juvenile Delinquency in Child Welfare: Investigating Group Home Effects," *Children and Youth Services Review*, 30, 1088–1099.

Serfes, K. (2005): "Risk Sharing vs. Incentives: Contract Design Under Two-sided Heterogeneity," *Economics Letters*, 88, 343–349.

Shimer, R. and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369.

Slaugh, V., M. Akan, O. Kesten, and U. Unver (2015): "The Pennsylvania Adoption Exchange Improves Its Matching Process," *Boston College Working Papers in Economics 858*.

U.S. Department of Health and Human Services (2020): "The AFCARS Reports," Retrieved from https://www.acf.hhs.gov/sites/default/files/documents/cb/afcarsreport28.pdf.

Van Tassel, E. (1999): "Group Lending Under Asymmetric Information," *Journal of Development Economics*, 60, 3–25.

Villodas, M., A. Litrownik, R. Newton, and I. Davis (2016): "Long-Term Placement Trajectories of Children who were Maltreated and Entered the Child Welfare System at an Early Age: Consequences for Physical and Behavioral Well-Being," *Journal of Pedriatric Psycology*, 41, 46–54.

A Appendix: Complete Information

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

$$\max_{\left\{\left(\lambda^k(x),\tau^k(x)\right)_x\right\}_{k=1}^2} \left\{ \sum_x \pi^c\big(\theta_x\big) \frac{\sum_k \left[u(x,y_k) - \tau^k(x)\right]\lambda^k(x)g(y_k)}{\sum_k \lambda^k(x)g(y_k)} f(x) \right\}$$

subject to:

$$[\text{FC}]\ \tau^k(x) \geq 0 \text{ and } \lambda^k(x) \geq 0 \text{ for all } (k,x),\ \text{and} \sum_x \lambda^k(x) = 1 \text{ for all } k$$

$$[\text{PC}]\ \sum_x \left[\tau^k(x) - c(x,y_k)\right]\lambda^k(x)\pi^p(\theta_x) \geq 0 \text{ for all } k$$

where

$$\theta_x = \frac{\sum_k \lambda^k(x)g(y_k)}{f(x)}$$

denotes the market tightness for child type x. Recall that

$$\pi^p\big(\theta\big) = \frac{\pi^c\big(\theta\big)}{\theta},$$

Hence the objective function becomes

$$\max_{\left\{\left(\lambda^k(x),\tau^k(x)\right)_x\right\}_{k=1}^2} \left\{ \sum_x \pi^p\big(\theta_x\big) \sum_k \left[u(x,y_k) - \tau^k(x)\right]\lambda^k(x)g(y_k) \right\}$$

To ease the notation, let $\pi^p(\theta_x) \equiv \pi^p_x$, $u(x,y_k) \equiv u^k_x$, $\tau^k(x) \equiv \tau^k_x$, and similarly $c(x,y_k) = x^k_x$. Now, it is easy to verify that

$$\sum_x \pi^p_x \sum_k \left[u^k_x - \tau^k_x\right]\lambda^k_x g_k = \sum_x \pi^p_x \left(\left[u^1_x - \tau^1_x\right]\lambda^1_x g_1 + \left[u^2_x - \tau^2_x\right]\lambda^2_x g_2\right)$$
$$= \pi^p_1 \left(\left[u^1_1 - \tau^1_1\right]\lambda^1_1 g_1 + \left[u^2_1 - \tau^2_1\right]\lambda^2_1 g_2\right) + \pi^p_2\left(\left[u^1_2 - \tau^1_2\right]\lambda^1_2 g_1 + \left[u^2_2 - \tau^2_2\right]\lambda^2_2 g_2\right)$$
$$= \left(\pi^p_1 u^1_1\lambda^1_1 + \pi^p_2 u^1_2\lambda^1_2 - \left[\pi^p_1\tau^1_1\lambda^1_1 + \pi^p_2\tau^1_2\lambda^1_2\right]\right)g_1 + \left(\pi^p_1 u^2_1\lambda^2_1 + \pi^p_2 u^2_2\lambda^2_2 - \left[\pi^p_1\tau^2_1\lambda^2_1 + \pi^p_2\tau^2_2\lambda^2_2\right]\right)g_2$$

Similarly, one can rewrite the participation constraints (which holds by equality) as the followings:

$$\text{PC}[1] : \tau_1^1 \lambda_1^1 \pi_1^p + \tau_2^1 \lambda_2^1 \pi_2^p = c_1^1 \lambda_1^1 \pi_1^p + c_2^1 \lambda_2^1 \pi_2^p,$$

and

$$\text{PC}[2] : \tau_1^2 \lambda_1^2 \pi_1^p + \tau_2^2 \lambda_2^2 \pi_2^p = c_1^2 \lambda_1^2 \pi_1^p + c_2^2 \lambda_2^2 \pi_2^p.$$

Therefore, plugging PC$[\cdot]$ into objective function, and rearranging it, the designer's problem becomes

$$\max_{\{\lambda_1^k, \lambda_2^k\}_{k=1}^2} \sum_x \pi_x^p \sum_k \left[u_x^k - c_x^k\right] \lambda_x^k g_k$$

subject to:

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

Notice the following corollary is immediate:

Corollary 4. *First best λ_x^y is independent from whether the designer is restricted with ex-post or intermediate participation constraints.*

Proof of Corollary 4. It simply follows from the fact that, the objective function becomes independent of transfers once the PCs are incorporated. This means ex-post and interim participation constraints does yield the same allocation of parents into submarkets. □

A.1 Proof of Lemma 1

To characterize the equilibrium λ_x^k, we follow a nonstandard technique for the following reason: there is always a corner solution for at least one λ_x^k and it is not obvious which one. Thus, the standard first order approach does not really yields us the solution. Denote $u_x^y - c_x^y \equiv U_x^y$ as the net benefit of the child-parent pair (x, y).

Let $\lambda_1^x \in (0, 1)$ be an arbitrary interior distribution of parents. Recall from Equation 3 that the welfare at that particular $(\lambda_1^1, \lambda_1^1)$:

$$W(\lambda_1^1, \lambda_1^2) = \pi_1^p \cdot \left(g_1 \lambda_1^1 U_1^1 + (1 - g_1) \lambda_1^2 U_1^2\right) + \pi_2^p \cdot \left(g_1 (1 - \lambda_1^1) U_2^1 + (1 - g_1)(1 - \lambda_1^2) U_2^2\right)$$

where

$$\theta_1 = \frac{g_1 \lambda_1^1 + (1 - g_1) \lambda_1^2}{f_1} \text{ and } \theta_2 = \frac{g_1 (1 - \lambda_1^1) + (1 - g_1)(1 - \lambda_1^2)}{1 - f_1}$$

Now change λ_1^1 by ε_1 and λ_1^2 by ε_2 such that $\varepsilon_2 \equiv -\frac{\varepsilon_1 g_1}{1-g_1}$, and thus, the market tightness θ_x for $x = 1, 2$ remain the same. The new welfare becomes

$$W(\lambda_1^1+\varepsilon_1, \lambda_1^2+\varepsilon_2) = \pi_1^p \cdot \left(g_1\lambda_1^1 U_1^1 + (1-g_1)\lambda_1^2 U_1^2\right) + \pi_2^p \cdot \left(g_1(1-\lambda_1^1)U_2^1 + (1-g_1)(1-\lambda_1^2)U_2^2\right)$$

$$\pi_1^p \varepsilon_1 g_1 U_1^1 - \pi_2^p \varepsilon_1 g_1 U_2^1 + \pi_1^p \varepsilon_2 (1-g_1) U_1^2 - \pi_2^p \varepsilon_2 (1-g_1) U_2^2.$$

By plugging $\varepsilon_2 = -\frac{\varepsilon_1 g_1}{1-g_1}$, the change in the welfare $\Delta_W = W(\lambda_1^1+\varepsilon_1, \lambda_1^2+\varepsilon_2) - W(\lambda_1^1, \lambda_1^2)$ is then

$$\Delta_W = \varepsilon_1 g_1 \underbrace{\left(\pi_2^p[U_2^2 - U_2^1] - \pi_1^p[U_1^2 - U_1^1]\right)}_{X(\theta)} \text{ where } \pi_x^p \equiv \pi^p(\theta_x) \text{for } x = 1, 2 \text{ and}$$

$$\theta_1 = \frac{g_1\lambda_1^1 + (1-g_1)\lambda_1^2}{f_1} \text{ and } \theta_2 = \frac{g_1(1-\lambda_1^1) + (1-g_1)(1-\lambda_1^2)}{1-f_1}.$$

It is easy to see that $X(\theta)$ is strictly increasing in θ_1. Therefore $X(\theta_1^{\max}) \geq X(\theta_1) \geq X(0)$ for any $\theta \in [0, \theta_1^{\max}]$ where $\theta_1^{\max} = \frac{1}{f_1}$ Notice, there are three different cases, i.e. $X(\theta) > 0$, $X(\theta) < 0$, and $X(\theta) = 0$.

1. Suppose $X(\theta) > 0$. Then pick $\varepsilon_1 > 0$ and $\varepsilon_2 < 0$ such that either $\lambda_1^1 = 1$ or $\lambda_1^2 = 0$, so that market tightness remains the same. This implies that the *forces* are in the PAM direction.

2. Suppose $X(\theta) < 0$. Then pick $\varepsilon_1 < 0$ and $\varepsilon_2 > 0$ such that either $\lambda_1^1 = 0$ or $\lambda_1^2 = 1$, so that market tightness remains the same. This implies that the *forces* are in the NAM direction.

3. Suppose $X(\theta) = 0$. Now, we show that this interior $\lambda_1^x \in (0, 1)$ cannot be an equilibrium. To see this, we first change λ_1^1 by ε_1 and evaluate the welfare, that is,

$$W_1 \equiv W(\lambda_1^1+\varepsilon_1, \lambda_1^2) = \hat{\pi}_1^p \cdot \left(g_1\lambda_1^1 U_1^1 + (1-g_1)\lambda_1^2 U_1^2\right) + \hat{\pi}_2^p \cdot \left(g_1(1-\lambda_1^1)U_2^1 + (1-g_1)(1-\lambda_1^2)U_2^2\right)$$

$$\hat{\pi}_1^p \varepsilon_1 g_1 U_1^1 - \hat{\pi}_2^p \varepsilon_1 g_1 U_2^1$$

$$\text{where } \hat{\pi}_2^p \equiv \pi^p(\hat{\theta}_2), \hat{\pi}_1^p \equiv \pi^p(\hat{\theta}_1), \text{and } \hat{\theta}_1 = \theta_1 + \frac{\varepsilon_1 g_1}{f_1}, \; \hat{\theta}_2 = \theta_2 - \frac{\varepsilon_1 g_1}{1-f_1}.$$

Now instead of λ_1^1, let's tremble λ_1^2 by ε_2 and evaluate the welfare, that is,

$$W_2 \equiv W(\lambda_1^1, \lambda_1^2{+}\varepsilon_2) = \tilde{\pi}_1^p \cdot \big(g_1\lambda_1^1 U_1^1 + (1-g_1)\lambda_1^2 U_1^2\big) + \tilde{\pi}_2^p \cdot \big(g_1(1-\lambda_1^1)U_2^1 + (1-g_1)(1-\lambda_1^2)U_2^2\big)$$

$$\tilde{\pi}_1^p \varepsilon_2 (1-g_1) U_1^2 - \tilde{\pi}_2^p \varepsilon_2 (1-g_1) U_2^2$$

where $\tilde{\pi}_2^p \equiv \pi^p(\tilde{\theta}_2), \tilde{\pi}_1^p \equiv \pi^p(\tilde{\theta}_1)$, and $\tilde{\theta}_1 = \theta_1 + \dfrac{\varepsilon_2(1-g_1)}{f_1}$, $\tilde{\theta}_2 = \theta_2 - \dfrac{\varepsilon_2(1-g_1)}{1-f_1}$.

Notice for any small ε_1, $\varepsilon_2 \equiv \frac{\varepsilon_1 g_1}{1-g_1}$ yields that $\tilde{\theta}_x = \hat{\theta}_x$ for $x = 1, 2$. Pick such an $\varepsilon_2(\varepsilon_1)$. Therefore, increasing λ_1^1 is marginally more profitable than increasing λ_1^2 if and only if

$$W_1 \geq W_2 \iff \underbrace{\pi^p\Big(\theta_2 - \frac{\varepsilon_1 g_1}{1-f_1}\Big) \cdot \big(U_2^2 - U_2^1\big) - \pi^p\Big(\theta_1 + \frac{\varepsilon_1 g_1}{f_1}\Big) \cdot \big(U_1^2 - U_1^1\big)}_{X(\hat{\theta})} \geq 0$$

where

$$\theta_1 = \frac{g_1\lambda_1^1 + (1-g_1)\lambda_1^2}{f_1} \text{ and } \theta_2 = \frac{g_1(1-\lambda_1^1) + (1-g_1)(1-\lambda_1^2)}{1-f_1}.$$

Note that $X(\hat{\theta}) > X(\theta) = 0$ implies that increasing λ_1^1 is marginally more profitable than increasing λ_1^2. Therefore, at least one of the *partial* derivatives of $W(\cdot,\cdot)$ at $(\lambda_1^1, \lambda_1^2)$ is non-zero, meaning that $(\lambda_1^1, \lambda_1^2)$ where $X(\theta) = 0$ is not an equilibrium. This finishes the proof.

A.2 Proof of Proposition 1

By construction U_x^y is increasing in y, and thus $X(\theta_1)$ is increasing in θ_1. Therefore, the three arguments about the sign of $x(\theta_1)$ in proof of Lemma 1 applies here.

A.3 Proof of Corollary 1

Proof of Corollary 1. Notice that $\theta_1 \in [0, \frac{1}{f_1}]$ which implies that $\theta_2 \in [0, \frac{1}{1-f_1}]$. Therefore, if $X(\theta_1^{\min} = 0) \geq 0$, then $X(\theta_1) \geq 0$ for any feasible θ_1 leading to PAM at the optimum. Analogous argument proves the second part of the Corollary. □

One can write Corollary 1 in a slightly different way. to do so, let $R_{ij} \equiv \frac{U_i^2 - U_i^1}{U_j^2 - U_j^1}$ denote the ratio of net utility differences between high-type and low-type parents i

across sub-markets x_i and x_j. The following is an alternative to Corollary 1:

Corollary 5. ***(i)*** *For any super-modular U_x^y, there exists $\underline{f} = 1 - \frac{1}{(\pi^p)^{-1}(R_{12})}$ such that the equilibrium exhibits PAM if $f_1 \le \max\{0, \underline{f}\}$.* ***(ii)*** *For any sup-modular U_x^y, there exists $\overline{f} = \frac{1}{(\pi^p)^{-1}(R_{21})}$ such that the equilibrium exhibits PAM if $f_1 \ge \min\{1, \overline{f}\}$.*

Proof of Corollary 5. It simply follows from Corollary 1 as well as the properties of meeting technology for parents $\pi^p(\cdot)$. □

Corollary 5 necessarily states that, low proportions of children with disability given super-modular net utility function imply PAM as the optimal sorting. Analogously, when the proportion of children with disability is high and the net utility is sub-modular, the optimal sorting becomes NAM. The rough intuition is as follows: High-ability parents are more desirable in any sub-market, the designer would like to allocate them to the thicker–more populated—market.

A.4 Characterization for the Equilibrium Allocation of Parents

Under the light of results above, we can characterize the equilibrium sorting patterns. Recall $W(\lambda_1^1, \lambda_1^2)$ defined as in Equation 3. First, we will prove a couple of auxiliary lemmas for the characterization of equilibrium parent allocations.

Lemma 3. *The rate of change in Welfare $W(\lambda_1^1, \lambda_1^2)$ monotonically decreases in λ_1^y for each $y = 1, 2$.*

Proof of Lemma 3. Recall the welfare of children:

$$W(\lambda_1^1, \lambda_1^2) = \pi^p(\theta_1) \cdot \underbrace{\left(g_1 \lambda_1^1 U_1^1 + (1-g_1)\lambda_1^2 U_1^2\right)}_{\mathbb{E}U_1} + \pi^p(\theta_2) \cdot \underbrace{\left(g_1(1-\lambda_1^1)U_2^1 + (1-g_1)(1-\lambda_1^2)U_2^2\right)}_{\mathbb{E}U_2}$$

where

$$\theta_1 = \frac{g_1 \lambda_1^1 + (1-g_1)\lambda_1^2}{f_1} \quad \text{and} \quad \theta_2 = \frac{1-\theta_1 f_1}{1-f_1}$$

Fix λ_1^{-y}. Increasing λ_1^y by a small amount $\varepsilon > 0$ increases $\mathbb{E}U_1$ and θ_1 while decreasing $\mathbb{E}U_2$ and θ_2 linearly. Recall $\pi^p(\cdot)$ is a decreasing, convex function which means the rate of increase through $\pi^p(\theta_1) \cdot \mathbb{E}U_1$ decreases whereas the rate of decrease through $\pi^p(\theta_2) \cdot \mathbb{E}U_2$ increases in λ_1^y, for any $y = 1, 2$. □

Lemma 3 is useful in proving the equilibrium sorting. It implies that $\frac{\partial W(\lambda_1^1,\lambda_1^2)}{\partial \lambda_1^y}$ is monotonically decreasing, and thus if it is zero at some λ_1^{y*}, then it is negative at any λ_1^y if and only if $\lambda_1^y > \lambda_1^{y*}$ for any λ_1^{-y}.

Note that exact same analysis applies to any pair of $(\lambda_1^1, \lambda_1^2)$ which yields the same market tightness θ_1, θ_2. Therefore another useful lemma follows:

Lemma 4. *Fix* $(\hat{\lambda}_1^1, \hat{\lambda}_1^2)$*. For any* $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2)$ *such that* $\theta_1(\hat{\lambda}_1^1, \hat{\lambda}_1^2) = \theta_1(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2)$ *and* $\hat{\lambda}_1^1 \geq \tilde{\lambda}_1^1$*,*

$$\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^y}|_{(\hat{\lambda}_1^1, \hat{\lambda}_1^2)} \leq \frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^y}|_{(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2)}.$$

Proof of Lemma 4. Taking the derivative of the welfare with respect to λ_1^y yields the followings:

$$W_1(\lambda_1^1, \lambda_1^2) \equiv \frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1} = g_1 \cdot Z(\lambda_1^1, \lambda_1^2) + g_1 \cdot \left(\pi^p(\theta_1)U_1^1 - \pi^p(\theta_2)U_2^1\right)$$

$$W_2(\lambda_1^1, \lambda_1^2) \equiv \frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^2} = (1-g_1) \cdot Z(\lambda_1^1, \lambda_1^2) + (1-g_1) \cdot \left(\pi^p(\theta_1)U_1^2 - \pi^p(\theta_2)U_2^2\right) \text{ where}$$

$$Z(\lambda_1^1, \lambda_1^2) = \frac{\pi^{p\prime}(\theta_1)}{f_1} \cdot \mathbb{E}U_1(\lambda_1^1, \lambda_1^2) - \frac{\pi^{p\prime}(\theta_2)}{1-f_1} \cdot \mathbb{E}U_2(\lambda_1^1, \lambda_1^2),$$

and $\mathbb{E}U_x(\lambda_1^1, \lambda_1^2)$ is defined as in Lemma 3. It is easy to verify that $Z(\lambda_1^1, \lambda_1^2)$ decreases as we move down on the market tightness θ_1, that is, as we increase λ_1^1 while decreasing λ_1^2. It implies that the rate of change with respect to λ_1^1 decreases as one moves down on the same market tightness, which finishes the proof. □

Now, by using Lemma 3 and Lemma 4, we characterize the equilibrium distribution of parents into submarket step by step. Initially, we establish the equilibrium allocation of parents whenever it is PAM or NAM described in Corollary 1. Later we extend the analysis where sufficient conditions in Corollary 1 are violated.

Proposition 5. *Suppose that* $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2 - U_1^1}{U_2^2 - U_2^1}$*. The equilibrium exhibits*

i. low-type partial PAM with $(\lambda_1^{1*}, 0)$ *if*

$$\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^2=0\}} = 0 \text{ for some } \lambda_1^{1*} \in (0,1), \tag{A.1}$$

ii. perfect PAM with $(1, 0)$ *if*

$$\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^1=1,\lambda_1^2=0\}} \geq 0 \geq \frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^2}|_{\{\lambda_1^1=1,\lambda_1^2=0\}}, \tag{A.2}$$

iii. high-type partial PAM with $(1, \lambda_1^{2*})$ *otherwise, where*

$$\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^2}|_{\{\lambda_1^1=1\}} = 0 \text{ for some } \lambda_1^{2*} \in (0, 1) \tag{A.3}$$

Proof of Proposition 5. It follows from the fact that $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2-U_1^1}{U_2^2-U_2^1}$ implies $X(\theta) \geq 0$ for any θ. Therefore, starting from the initial allocation $\lambda_1^y = 0$ for each $y = 1, 2$, the designer starts allocating type-1 parents into submarket-1 until it is no more profitable to do so or they are exhausted. Accordingly, perfect PAM and high-type partial PAM follows. One can easily characterize the equilibrium distribution of parents when it is always NAM, with a parallel argument. □

One can easily characterize the equilibrium distribution of parents when it is always NAM, with a parallel argument.

Proposition 6. *Suppose that* $\pi^p\left(\frac{1}{f_1}\right) \geq \frac{U_2^2-U_2^1}{U_1^2-U_1^1}$. *The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ *if*

$$\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^2}|_{\{\lambda_1^1=0\}} = 0 \text{ for some } \lambda_1^{2*} \in (0, 1), \tag{A.4}$$

ii. perfect NAM with $(0, 1)$ *if*

$$\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^2}|_{\{\lambda_1^1=0,\lambda_1^2=1\}} \geq 0 \geq \frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^1=0,\lambda_1^2=1\}}, \tag{A.5}$$

iii. low-type partial NAM with $(\lambda_1^{1*}, 1)$ *otherwise, where*

$$\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^2=1\}} = 0 \text{ for some } \lambda_1^{1*} \in (0, 1). \tag{A.6}$$

Proof of Proposition 6. QED following the same arguments in the proof of Proposition 5 □

Propositions 5 and 6 characterize the equilibrium distribution of parents when the environment yields monotonic sorting—either PAM or NAM—irrespective of other parameters. One wonders how to characterize the equilibrium distribution of parents into submarkets when the equilibrium sorting is contingent upon parameters. To study such phenomenon, we suppose that $X(\theta_1) = 0$ for some $\theta_1 \in (0, \frac{1}{f_1})$. Then there exists either $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ with $\tilde{\lambda}_1^2 \leq 1$ or $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ with $\tilde{\lambda}_1^1 > 0$ such that $\theta_1 = \frac{g_1\tilde{\lambda}_1^1+(1-g_1)\tilde{\lambda}_1^2}{f_1}$. Similarly, there also exists either $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$ or $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 > 0$ such that $\theta_1 = \frac{g_1\hat{\lambda}_1^1+(1-g_1)\hat{\lambda}_1^2}{f_1}$ (See top 4 charts in Figure 3). In what follows, we study each of these cases separately.

Case 1. Suppose $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ with $\tilde{\lambda}_1^2 \leq 1$; i.e. we are at a situation as in the top two chart of Figure 3. Now, we consider $\hat{\lambda}_1^1$ as sub-cases

i. Suppose $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$, i.e. we are at a situation as in the top left chart of Figure 3. Starting from a particular corner $(\lambda_1^1 = 0, \lambda_1^2 = 0)$, we initially increase λ_1^2, until it is not profitable any more to increase λ_1^2, or we switch from NAM to PAM and increase λ_1^1 instead. The following characterizes equilibrium sorting for this particular case

Proposition 7. *Suppose* $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ *with* $\tilde{\lambda}_1^2 \leq 1$ *and* $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ *with* $\hat{\lambda}_1^1 \leq 1$. *The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ *if* $\lambda_1^{2*} < \tilde{\lambda}_1^2$ *where* λ_1^{2*} *solves (A.4).*

ii. Suppose $\lambda_1^{2*} \geq \tilde{\lambda}_1^2$. *Then the equilibrium exhibits*

a. low-type partial PAM with $(\lambda_1^{1*}, 0)$ *with* $(\lambda_1^{1*}, 0)$, *if where* $\lambda_1^{1*} \in (\hat{\lambda}_1^1, 1)$ *where* λ_1^{1*} *solves (A.1)*

b. perfect PAM if (A.2) holds, and

c. high-type partial PAM with $(1, \lambda_1^{2*})$ *otherwise, where* λ_1^{2*} *solves (A.3)*

ii. Suppose $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 \geq 0$, i.e. we are at a situation as in top right chart of Figure 3. Following a symmetric argument, the characterization of this case follows:

Proposition 8. *Suppose* $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ *with* $\tilde{\lambda}_1^2 \leq 1$ *and* $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ *with* $\hat{\lambda}_1^2 \geq 0$. *The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ *if* $\lambda_1^{2*} < \tilde{\lambda}_1^2$ *where* λ_1^{2*} *solves (A.4).*

ii. high-type partial PAM with $(1, \lambda_1^{2*})$ *otherwise, where* λ_1^{2*} *solves (A.3).*

Case 2. Suppose $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ with $\tilde{\lambda}_1^1 > 0$; i.e. we are at a situation as in the middle two chart of Figure 3. Now, we consider the sub-cases of $\hat{\lambda_1^1}$ as in the analyses above

i. Suppose $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$, i.e. we are at a situation as in the middle-left chart of Figure 3—which requires the slope to be less than -1. The following characterizes equilibrium sorting in this case:

Proposition 9. *Suppose $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ with $\tilde{\lambda}_1^1 > 0$ and $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$. The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ if $\lambda_1^{2*} < \tilde{\lambda}_1^2$ where λ_1^{2*} solves (A.4).

ii. Suppose $\lambda_1^{2*} \geq \tilde{\lambda}_1^2$. Then the equilibrium exhibits

a. perfect NAM if $\frac{\partial W(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^1=0, \lambda_1^2=1\}} \leq 0$

b. low-type partial NAM with $(\lambda_1^{1*}, 1)$, if $\lambda_1^{1*} \leq \tilde{\lambda}_1^1$ where λ_1^{1*} solves (A.6)

c. Suppose $\lambda_1^{1*} > \tilde{\lambda}_1^2$ where λ_1^{1*} solves (A.6). Then the equilibrium exhibits

[1.] low-type partial PAM with $(\lambda_1^{1*}, 0)$, if $\lambda_1^{1*} < 1$ where λ_1^{1*} solves (A.1)

[2.] perfect PAM if (A.2) holds, and

[3.] high-type partial PAM with $(1, \lambda_1^{2*})$, otherwise, where λ_1^{2*} solves (A.3)

ii. Suppose $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 \geq 0$, i.e., we are at a situation as in middle-right chart of Figure 3. Following a symmetric argument, the characterization of this case follows:

Proposition 10. *Suppose $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ with $\tilde{\lambda}_1^1 > 0$ and$(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 > 0$. The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ if $\lambda_1^{2*} < 1$ where λ_1^{2*} solves (A.4).

ii. perfect NAM if (A.5) holds, and

iii. low-type partial NAM with $(\lambda_1^{1*}, 1)$ where λ_1^{1*} solves (A.6), otherwise.

B Appendix: Private Information

In this section, we analyze the optimal licenses for an environment where the designer does not observe the parent characteristics, while the child characteristic is observable. This extension is motivated by the common phenomenon in which the ability of a foster parent is a private information to him/herself. First it is useful to understand who has incentives to mimic whom under the first best menu of licenses. To ease the notation, let $\pi^p(\theta_{x_i}) \equiv \pi_i^p$, $u(x_i, y_k) \equiv u_i^k$, $\tau^k(x_i) \equiv \tau_i^k$, and similarly $c(x_i, y_k) = c_i^k$. Recall the incentive compatibility constraint $IC[i]$ for $i \neq j \in \{1, 2\}$,

$$\sum_x \left[\tau^i(x) - c(x, y_i)\right] \lambda^i(x) \pi^p(\theta(x)) \geq \sum_x \left[\tau^j(x) - c(x, y_i)\right] \lambda^j(x) \pi^p(\theta(x)) \text{ iff}$$
$$\tau_1^i \lambda_1^i \pi_1^p + \tau_2^i \lambda_2^i \pi_2^p - c_1^i \lambda_1^i \pi_1^p - c_2^i \lambda_2^i \pi_2^p \geq \tau_1^j \lambda_1^j \pi_1^p + \tau_2^j \lambda_2^j \pi_2^p - c_1^i \lambda_1^j \pi_1^p - c_2^i \lambda_2^j \pi_2^p$$

Recall also that the participation constraint $PC[i]$ for $i = 1, 2$:

$$\tau_1^i \lambda_1^i \pi_1^p + \tau_2^i \lambda_2^i \pi_2^p = c_1^i \lambda_1^i \pi_1^p + c_2^i \lambda_2^i \pi_2^p.$$

Thus plugging $PC[i]$ and $PC[j]$ into $IC[i]$ yields:

$$IC[i]: \; c_1^i \lambda_1^i \pi_1^p + c_2^i \lambda_2^i \pi_2^p - c_1^i \lambda_1^i \pi_1^p - c_2^i \lambda_2^i \pi_2^p \geq c_1^j \lambda_1^j \pi_1^p + c_2^j \lambda_2^j \pi_2^p - c_1^i \lambda_1^j \pi_1^p - c_2^i \lambda_2^j \pi_2^p \quad \text{iff}$$
$$0 \geq (c_1^j - c_1^i) \lambda_1^j \pi_1^p + (c_2^j - c_2^i) \lambda_2^j \pi_2^p$$

Recall that c_x^k is decreasing in both components, and hence the inequality holds for $i = 1$ but not for $i = 2$, for any feasible λ_x^k. This means, under the first best, type-1 parent does not have incentives to mimic type-2 parent but we have the mimic of the other way around. Therefore, $IC[2]$ and $PC[1]$ hold by equality in equilibrium:

$$\text{PC}[1]: \tau_1^1 \lambda_1^1 \pi_1^p + \tau_2^1 \lambda_2^1 \pi_2^p = c_1^1 \lambda_1^1 \pi_1^p + c_2^1 \lambda_2^1 \pi_2^p, \quad \text{and}$$

$$\text{IC}[2]: \tau_1^2 \lambda_1^2 \pi_1^p + \tau_2^2 \lambda_2^2 \pi_2^p - c_1^2 \lambda_1^2 \pi_1^p - c_2^2 \lambda_2^2 \pi_2^p = \underbrace{\tau_1^1 \lambda_1^1 \pi_1^p + \tau_2^1 \lambda_2^1 \pi_2^p}_{c_1^1 \lambda_1^1 \pi_1^p + c_2^1 \lambda_2^1 \pi_2^p} - c_1^2 \lambda_1^1 \pi_1^p - c_2^2 \lambda_2^1 \pi_2^p$$

which then becomes the following

$$\tau_1^2 \lambda_1^2 \pi_1^p + \tau_2^2 \lambda_2^2 \pi_2^p = c_1^2 \lambda_1^2 \pi_1^p + c_2^2 \lambda_2^2 \pi_2^p + (c_1^1 - c_1^2) \lambda_1^1 \pi_1^p + (c_2^1 - c_2^2) \lambda_2^1 \pi_2^p$$

Recall the objective function:

$$\max_{\left\{\left(\lambda^k(x),\tau^k(x)\right)_x\right\}_{k=1}^2} \left\{\sum_x \pi^p\left(\theta_x\right) \sum_k \left[u(x,y_k)-\tau^k(x)\right]\lambda^k(x)g(y_k)\right\}$$

Now, it is easy to verify that

$$\begin{aligned}
&\sum_x \pi_x^p \sum_k \left[u_x^k-\tau_x^k\right]\lambda_x^k g_k = \sum_x \pi_x^p \left(\left[u_x^1-\tau_x^1\right]\lambda_x^1 g_1 + \left[u_x^2-\tau_x^2\right]\lambda_x^2 g_2\right)\\
&= \pi_1^p \left(\left[u_1^1-\tau_1^1\right]\lambda_1^1 g_1 + \left[u_1^2-\tau_1^2\right]\lambda_1^2 g_2\right) + \pi_2^p\left(\left[u_2^1-\tau_2^1\right]\lambda_2^1 g_1 + \left[u_2^2-\tau_2^2\right]\lambda_2^2 g_2\right)\\
&= \left(\pi_1^p u_1^1\lambda_1^1 + \pi_2^p u_2^1\lambda_2^1 - \left[\pi_1^p\tau_1^1\lambda_1^1 + \pi_2^p\tau_2^1\lambda_2^1\right]\right)g_1 + \left(\pi_1^p u_1^2\lambda_1^2 + \pi_2^p u_2^2\lambda_2^2 - \left[\pi_1^p\tau_1^2\lambda_1^2 + \pi_2^p\tau_2^2\lambda_2^2\right]\right)g_2
\end{aligned}$$

Plugging $PC[1]$ and $IC[2]$ into the objective function yields

$$\begin{aligned}
&\sum_x \pi_x^p \sum_k \left[u_x^k-\tau_x^k\right]\lambda_x^k g_k =\\
&= \left(\pi_1^p u_1^1\lambda_1^1 + \pi_2^p u_2^1\lambda_2^1 - \left[c_1^1\lambda_1^1\pi_1^p + c_2^1\lambda_2^1\pi_2^p\right]\right)g_1 +\\
&+\left(\pi_1^p u_1^2\lambda_1^2 + \pi_2^p u_2^2\lambda_2^2 - \left[c_1^2\lambda_1^2\pi_1^p + c_2^2\lambda_2^2\pi_2^p + (c_1^1-c_1^2)\lambda_1^1\pi_1^p + (c_2^1-c_2^2)\lambda_2^1\pi_2^p\right]\right)g_2 =\\
&= \left(\pi_1^p\lambda_1^1[u_1^1-c_1^1] + \pi_2^p\lambda_2^1[u_2^1-c_2^1]\right)g_1 +\\
&+\left(\pi_1^p\lambda_1^2[u_1^2-c_1^2] + \pi_2^p\lambda_2^2[u_2^2-c_2^2] - (c_1^1-c_1^2)\lambda_1^1\pi_1^p - (c_2^1-c_2^2)\lambda_2^1\pi_2^p\right)g_2
\end{aligned}$$

Recall the definition $U_x^k = u_x^k - c_x^k$, and denote $\Delta c_x = (c_x^1 - c_x^2) > 0$, the objective function becomes

$$\max_{\left\{\lambda_1^1,\lambda_1^2\right\}} \left\{\left(\pi_1^p\lambda_1^1U_1^1 + \pi_2^p(1-\lambda_1^1)U_2^1\right)g_1 + \left(\pi_1^p\lambda_1^2U_1^2 + \pi_2^p(1-\lambda_1^2)U_2^2 - \Delta c_1\lambda_1^1\pi_1^p - \Delta c_2(1-\lambda_1^1)\pi_2^p\right)g_2\right\}$$

Now, we can establish the corner solution result under the presence of information frictions, as in the complete information case:

Proof of Lemma 2. Let $\lambda_1^x \in (0,1)$ be an arbitrary interior distribution of parents. Write the welfare at that particular $(\lambda_1^1, \lambda_1^2)$:

$$\hat{W}(\lambda_1^1, \lambda_1^2) = \Big(\pi_1^p \lambda_1^1 U_1^1 + \pi_2^p (1-\lambda_1^1) U_2^1\Big) g_1 + \Big(\pi_1^p \lambda_1^2 U_1^2 + \pi_2^p (1-\lambda_1^2) U_2^2 - \Delta c_1 \lambda_1^1 \pi_1^p - \Delta c_2 (1-\lambda_1^1) \pi_2^p\Big) g_2 \tag{B.1}$$

As in the complete information, we change λ_1^1 by ε_1 and λ_1^2 by ε_2 such that $\varepsilon_2 \equiv -\frac{\varepsilon_1 g_1}{1-g_1}$, and thus, the market tightness θ_x for $x = 1, 2$ remain the same. The new welfare becomes

$$\hat{W}(\lambda_1^1+\varepsilon_1, \lambda_1^2+\varepsilon_2) = \Big(\pi_1^p \lambda_1^1 U_1^1 + \pi_2^p (1-\lambda_1^1) U_2^1\Big) g_1 + \Big(\pi_1^p \lambda_1^2 U_1^2 + \pi_2^p (1-\lambda_1^2) U_2^2 - \Delta c_1 \lambda_1^1 \pi_1^p - \Delta c_2 (1-\lambda_1^1) \pi_2^p\Big) g_2$$

$$\pi_1^p \varepsilon_1 g_1 U_1^1 - \pi_2^p \varepsilon_1 g_1 U_2^1 + \pi_1^p \varepsilon_2 g_2 U_1^2 - \pi_2^p \varepsilon_2 g_2 U_2^2 - \pi_1^p \varepsilon_1 g_2 \Delta c_1 + \pi_2^p \varepsilon_1 g_2 \Delta c_2$$

By plugging $\varepsilon_2 = -\frac{\varepsilon_1 g_1}{1-g_1}$, the change in the welfare $\Delta_{\hat{W}} = \hat{W}(\lambda_1^1+\varepsilon_1, \lambda_1^2+\varepsilon_2) - \hat{W}(\lambda_1^1, \lambda_1^2)$ is then

$$\Delta_{\hat{W}} = \varepsilon_1 g_1 \underbrace{\left[\pi_2^p \cdot \left(U_2^2 - U_2^1 + \frac{g_2}{g_1}\Delta c_2\right) - \pi_1^p \cdot \left(U_1^2 - U_1^1 + \frac{g_2}{g_1}\Delta c_1\right)\right]}_{\hat{X}(\theta)} \text{ where } \pi_x^p \equiv \pi^p(\theta_x) \text{for } x = 1, 2 \text{ and}$$

$$\theta_1 = \frac{g_1 \lambda_1^1 + (1-g_1)\lambda_1^2}{f_1} \;\text{ and }\; \theta_2 = \frac{g_1(1-\lambda_1^1) + (1-g_1)(1-\lambda_1^2)}{1-f_1}.$$

As in earlier, $\hat{X}(\theta)$ is strictly increasing in θ_1. Therefore $\hat{X}(\theta_1^{\max}) \geq \hat{X}(\theta_1) \geq \hat{X}(0)$ for any $\theta \in [0, \theta_1^{\max}]$ where $\theta_1^{\max} = \frac{1}{f_1}$ Notice, there are three different cases, i.e. $\hat{X}(\theta) > 0$, $\hat{X}(\theta) < 0$, and $\hat{X}(\theta) = 0$.

1. Suppose $\hat{X}(\theta) > 0$. Then pick $\varepsilon_1 > 0$ and $\varepsilon_2 < 0$ such that either $\lambda_1^1 = 1$ or $\lambda_1^2 = 0$, so that market tightness remains the same. This implies that the *forces* are in the PAM direction. Since $\hat{X}(\theta)$ is also increasing in θ_1, it is easy to establish a sufficient condition for PAM:

   Lemma 5. *If* $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2 - U_1^1 + \frac{g_2}{g_1}\Delta c_1}{U_2^2 - U_2^1 + \frac{g_2}{g_1}\Delta c_2}$ *then the equilibrium exhibits (perfect or partial) PAM.*

2. Suppose $\hat{X}(\theta) < 0$. Then pick $\varepsilon_1 < 0$ and $\varepsilon_2 > 0$ such that either $\lambda_1^1 = 0$ or $\lambda_1^2 = 1$, so that market tightness remains the same. This implies that the *forces* are in the NAM direction. An analogous sufficient condition for NAM exists:

Lemma 6. *If* $\pi^p\Big(\frac{1}{f_1}\Big) \geq \frac{U_2^2-U_2^1+\frac{g_2}{g_1}\Delta c_2}{U_1^2-U_1^1+\frac{g_2}{g_1}\Delta c_1}$ *then the equilibrium exhibits* (*perfect or partial*) *NAM.*

Remark 1. *Note as in the complete information setting, supermodularity* (*submodularity*) *of* U_x^y *is not sufficient for PAM* (*NAM*)*, and even further, it is not necessarily the case even if there were no search frictions— that is, if* $\pi^p(\cdot) = 1$*.*

3. Suppose $\hat{X}(\theta) = 0$. Now, we will show that this interior $\lambda_1^x \in (0,1)$ cannot be an equilibrium. To see this, we will first change λ_1^1 by ε_1 and evaluate the welfare, that is, $\hat{W}_1 \equiv \hat{W}(\lambda_1^1+\varepsilon_1, \lambda_1^2)$

$$\hat{W}_1 = \Big(\hat{\pi}_1^p\lambda_1^1U_1^1+\hat{\pi}_2^p(1-\lambda_1^1)U_2^1\Big)g_1+\Big(\hat{\pi}_1^p\lambda_1^2U_1^2+\hat{\pi}_2^p(1-\lambda_1^2)U_2^2-\Delta c_1\lambda_1^1\hat{\pi}_1^p-\Delta c_2(1-\lambda_1^1)\hat{\pi}_2^p\Big)g_2$$

$$\hat{\pi}_1^p\varepsilon_1g_1U_1^1 - \hat{\pi}_2^p\varepsilon_1g_1U_2^1 - \hat{\pi}_1^p\varepsilon_1g_2\Delta c_1 + \hat{\pi}_2^p\varepsilon_1g_2\Delta c_2$$

$$\text{where} \quad \hat{\pi}_2^p \equiv \pi^p(\hat{\theta}_2), \hat{\pi}_1^p \equiv \pi^p(\hat{\theta}_1), \text{and } \hat{\theta}_1 = \theta_1 + \frac{\varepsilon_1 g_1}{f_1},\ \hat{\theta}_2 = \theta_2 - \frac{\varepsilon_1 g_1}{1-f_1}.$$

Now instead, change λ_1^2 by ε_2 and evaluate the welfare, that is, $\hat{W}_2 \equiv \hat{W}(\lambda_1^1, \lambda_1^2 + \varepsilon_2)$

$$\hat{W}_2 = \Big(\tilde{\pi}_1^p\lambda_1^1U_1^1+\tilde{\pi}_2^p(1-\lambda_1^1)U_2^1\Big)g_1+\Big(\tilde{\pi}_1^p\lambda_1^2U_1^2+\tilde{\pi}_2^p(1-\lambda_1^2)U_2^2-\Delta c_1\lambda_1^1\tilde{\pi}_1^p-\Delta c_2(1-\lambda_1^1)\tilde{\pi}_2^p\Big)g_2$$

$$+\tilde{\pi}_1^p\varepsilon_2g_2U_1^2 - \tilde{\pi}_2^p\varepsilon_2g_2U_2^2$$

$$\text{where} \quad \tilde{\pi}_2^p \equiv \pi^p(\tilde{\theta}_2), \tilde{\pi}_1^p \equiv \pi^p(\tilde{\theta}_1), \text{and } \tilde{\theta}_1 = \theta_1 + \frac{\varepsilon_2(1-g_1)}{f_1},\ \tilde{\theta}_2 = \theta_2 - \frac{\varepsilon_2(1-g_1)}{1-f_1}.$$

Notice for any small ε_1, $\varepsilon_2 \equiv \frac{\varepsilon_1 g_1}{1-g_1}$ yields that $\tilde{\theta}_x = \hat{\theta}_x$ for $x = 1, 2$. Pick such an $\varepsilon_2(\varepsilon_1)$. Therefore, increasing λ_1^1 is marginally more profitable than increasing λ_1^2 if and only if $\hat{W}_1 \geq \hat{W}_2$, that is

$$\underbrace{\pi^p\Big(\theta_2 - \frac{\varepsilon_1 g_1}{1-f_1}\Big)\cdot\big(U_2^2 - U_2^1 + \frac{g_2}{g_1}\Delta c_2\big) - \pi^p\Big(\theta_1 + \frac{\varepsilon_1 g_1}{f_1}\Big)\cdot\big(U_1^2 - U_1^1 + \frac{g_2}{g_1}\Delta c_1\big)}_{\hat{X}(\hat{\theta})} \geq 0$$

where

$$\theta_1 = \frac{g_1\lambda_1^1 + (1-g_1)\lambda_1^2}{f_1} \ \text{ and } \ \theta_2 = \frac{g_1(1-\lambda_1^1) + (1-g_1)(1-\lambda_1^2)}{1-f_1}.$$

Note that $\hat{X}(\hat{\theta}) > \hat{X}(\theta) = 0$ implies that increasing λ_1^1 is marginally more profitable than increasing λ_1^2. Therefore, at least one of the *partial* derivatives of $\hat{W}(\cdot,\cdot)$ at $(\lambda_1^1, \lambda_1^2)$ is non-zero, meaning that $(\lambda_1^1, \lambda_1^2)$ where $X(\theta) = 0$ is not an equilibrium. This finishes our proof.

□

B.1 Characterization for the Equilibrium Allocation of Parents

Before full characterization of the equilibrium sorting, we show that Lemma 3 and Lemma 4 carries over to the case of private asymmetry.

Lemma 7. *The rate of change in Welfare $\hat{W}(\lambda_1^1, \lambda_1^2)$ monotonically decreases in λ_1^y for each $y = 1, 2$.*

Proof of Lemma 7. Recall the welfare of children:

$$\begin{aligned}\hat{W}(\lambda_1^1, \lambda_1^2) &= \pi^p(\theta_1) \cdot \overbrace{\left(g_1 \lambda_1^1 U_1^1 + (\lambda_1^2 U_1^2 - \lambda_1^1 \Delta c_1) g_2\right)}^{\hat{\mathbb{E}}U_1} \\ &+ \pi^p(\theta_2) \cdot \underbrace{\Big(g_1(1-\lambda_1^1) U_2^1 + \big((1-\lambda_1^2) U_2^2 - (1-\lambda_1^1)\Delta c_2\big) g_2\Big)}_{\hat{\mathbb{E}}U_2}\end{aligned}$$

where

$$\theta_1 = \frac{g_1 \lambda_1^1 + (1-g_1)\lambda_1^2}{f_1} \text{ and } \theta_2 = \frac{1-\theta_1 f_1}{1-f_1}$$

Fixing λ_1^{-y}, increasing λ_1^y by a small amount $\varepsilon > 0$ increases $\hat{\mathbb{E}}U_1$ and θ_1 while decreasing $\hat{\mathbb{E}}U_2$ and θ_2 linearly. Recall $\pi^p(\cdot)$ is a decreasing, convex function which means the rate of increase through $\pi^p(\theta_1) \cdot \hat{\mathbb{E}}U_1$ decreases whereas the rate of decrease through $\pi^p(\theta_2) \cdot \hat{\mathbb{E}}U_2$ increases in λ_1^y, for any $y = 1, 2$. □

Lemma 8. *Fix $(\hat{\lambda}_1^1, \hat{\lambda}_1^2)$. For any $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2)$ such that $\theta_1(\hat{\lambda}_1^1, \hat{\lambda}_1^2) = \theta_1(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2)$ and $\hat{\lambda}_1^1 \geq \tilde{\lambda}_1^1$,*

$$\frac{\partial \hat{W}(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^y}\Big|_{(\hat{\lambda}_1^1, \hat{\lambda}_1^2)} \leq \frac{\partial \hat{W}(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^y}\Big|_{(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2)}.$$

Proof of Lemma 8. Taking the derivative of the welfare under private information with respect to λ_1^y yields the followings:

$$\hat{W}_1(\lambda_1^1, \lambda_1^2) \equiv \frac{\partial \hat{W}(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1} = g_1 \cdot Z(\lambda_1^1, \lambda_1^2) + g_1 \cdot \big(\pi^p(\theta_1) U_1^1 - \pi^p(\theta_2) U_2^1\big) - g_2 \cdot \big(\pi^p(\theta_1)\Delta c_1 - \pi^p(\theta_2)\Delta c_2\big)$$

$$\hat{W}_2(\lambda_1^1,\lambda_1^2) \equiv \frac{\partial \hat{W}(\lambda_1^1,\lambda_1^2)}{\partial \lambda_1^2} = g_2 \cdot Z(\lambda_1^1,\lambda_1^2) + g_2 \cdot \left(\pi^p(\theta_1)U_1^2 - \pi^p(\theta_2)U_2^2\right) \text{ where}$$

$$Z(\lambda_1^1,\lambda_1^2) = \frac{\pi^{p\prime}(\theta_1)}{f_1} \cdot \hat{\mathbb{E}}U_1(\lambda_1^1,\lambda_1^2) - \frac{\pi^{p\prime}(\theta_2)}{1-f_1} \cdot \hat{\mathbb{E}}U_2(\lambda_1^1,\lambda_1^2),$$

and $\hat{\mathbb{E}}U_x(\lambda_1^1,\lambda_1^2)$ is defined as in Lemma 7. Notice, plugging $Z(\lambda_1^1,\lambda_1^2)$ into $\hat{W}_1(\lambda_1^1,\lambda_1^2)$ yields the following:

$$\hat{W}_1(\lambda_1^1,\lambda_1^2) = \frac{g_1}{g_2} \cdot \hat{W}_2(\lambda_1^1,\lambda_1^2) + g_1 \cdot \hat{X}(\theta).$$

Note also that $Z(\lambda_1^1,\lambda_1^2)$ is the same as in complete information case, and thus it decreases as we move down on the market tightness θ_1. In other words, as we increase λ_1^1 while decreasing λ_1^2, $Z(\lambda_1^1,\lambda_1^2)$ decreases. It implies that the rate of change with respect to λ_1^1 decreases as one moves down on the same market tightness, which finishes the proof. □

Now, by using Lemma 7 and Lemma 8, we characterize the equilibrium distribution of parents into submarket as in the complete information case. Initially, we establish the equilibrium allocation of parents whenever it is PAM or NAM described in Corollary 2. Later we extend the analysis where sufficient conditions in Corollary 2 are violated.

Proposition 11. *Suppose that* $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2 - U_1^1 + \frac{g_2}{g_1}\Delta c_1}{U_2^2 - U_2^1 + \frac{g_2}{g_1}\Delta c_2}$. *The equilibrium exhibits*

i. low-type partial PAM with $(\lambda_1^{1*}, 0)$ *if*

$$\frac{\partial \hat{W}(\lambda_1^1,\lambda_1^2)}{\partial \lambda_1^1}\Big|_{\{\lambda_1^2=0\}} = 0 \text{ for some } \lambda_1^{1*} \in (0,1), \tag{B.2}$$

ii. perfect PAM with $(1,0)$ *if*

$$\frac{\partial \hat{W}(\lambda_1^1,\lambda_1^2)}{\partial \lambda_1^1}\Big|_{\{\lambda_1^1=1,\lambda_1^2=0\}} \geq 0 \geq \frac{\partial \hat{W}(\lambda_1^1,\lambda_1^2)}{\partial \lambda_1^2}\Big|_{\{\lambda_1^1=1,\lambda_1^2=0\}}, \tag{B.3}$$

iii. high-type partial PAM with $(1, \lambda_1^{2*})$ *otherwise, where*

$$\frac{\partial \hat{W}(\lambda_1^1,\lambda_1^2)}{\partial \lambda_1^2}\Big|_{\{\lambda_1^1=1\}} = 0 \text{ for some } \lambda_1^{2*} \in (0,1) \tag{B.4}$$

Proof of Proposition 11. It follows from the fact that $\pi^p\left(\frac{1}{1-f_1}\right) \geq \frac{U_1^2-U_1^1+\frac{g_2}{g_1}\Delta c_1}{U_2^2-U_2^1+\frac{g_2}{g_1}\Delta c_2}$ implies $X(\theta) \geq 0$ for any θ. Therefore, starting from the initial allocation $\lambda_1^y = 0$ for each $y = 1, 2$, the designer starts allocating type-1 parents into submarket-1 until it is no more profitable to do so or they are exhausted. Accordingly, perfect PAM and high-type partial PAM follows. One can easily characterize the equilibrium distribution of parents when it is always NAM, with a parallel argument. □

Analogously, we characterize the equilibrium distribution of parents when it is always NAM, with a parallel argument.

Proposition 12. *Suppose that* $\pi^p\left(\frac{1}{f_1}\right) \geq \frac{U_2^2-U_2^1+\frac{g_2}{g_1}\Delta c_2}{U_1^2-U_1^1+\frac{g_2}{g_1}\Delta c_1}$. *The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ *if*

$$\frac{\partial \hat{W}(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^2}|_{\{\lambda_1^1=0\}} = 0 \text{ for some } \lambda_1^{2*} \in (0,1), \tag{B.5}$$

ii. perfect NAM with $(0, 1)$ *if*

$$\frac{\partial \hat{W}(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^2}|_{\{\lambda_1^1=0,\lambda_1^2=1\}} \geq 0 \geq \frac{\partial \hat{W}(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^1=0,\lambda_1^2=1\}}, \tag{B.6}$$

iii. low-type partial NAM with $(\lambda_1^{1*}, 1)$ *otherwise, where*

$$\frac{\partial \hat{W}(\lambda_1^1, \lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^2=1\}} = 0 \text{ for some } \lambda_1^{1*} \in (0,1). \tag{B.7}$$

Proof of Proposition 12. QED following the same arguments in the proof of Proposition 5 □

Propositions 11 and 12 characterize the equilibrium distribution of parents when the environment yields monotonic sorting—either PAM or NAM—irrespective of other parameters. To further characterize the equilibrium allocation when conditions above are violated, we suppose that $X(\theta_1) = 0$ for some $\theta_1 \in (0, \frac{1}{f_1})$. Then there exists either $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ with $\tilde{\lambda}_1^2 \leq 1$ or $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ with $\tilde{\lambda}_1^1 > 0$ such that $\theta_1 = \frac{g_1\tilde{\lambda}_1^1+(1-g_1)\tilde{\lambda}_1^2}{f_1}$. Similarly, there also exists either $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$ or $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 > 0$ such that $\theta_1 = \frac{g_1\hat{\lambda}_1^1+(1-g_1)\hat{\lambda}_1^2}{f_1}$ (See top 4 charts in Figure 3). In what follows, we study each of these cases separately.

Case 1. Suppose $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ with $\tilde{\lambda}_1^2 \leq 1$; i.e. we are at a situation as in the top two chart of Figure 3. Now, we consider $\hat{\lambda}_1^1$ as sub-cases

i. Suppose $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$, i.e. we are at a situation as in the top left chart of Figure 3. Starting from a particular corner $(\lambda_1^1 = 0, \lambda_1^2 = 0)$, we initially increase λ_1^2, until it is not profitable any more to increase λ_1^2, or we switch from NAM to PAM and increase λ_1^1 instead. The following characterizes equilibrium sorting for this particular case

Proposition 13. *Suppose $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ with $\tilde{\lambda}_1^2 \leq 1$ and $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$. The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ if $\lambda_1^{2*} < \tilde{\lambda}_1^2$ where λ_1^{2*} solves (B.5).

ii. Suppose $\lambda_1^{2*} \geq \tilde{\lambda}_1^2$. Then the equilibrium exhibits

a. low-type partial PAM with $(\lambda_1^{1*}, 0)$ with $(\lambda_1^{1*}, 0)$, if where $\lambda_1^{1*} \in (\hat{\lambda}_1^1, 1)$ where λ_1^{1*} solves (B.2)

b. perfect PAM if (B.3) holds, and

c. high-type partial PAM with $(1, \lambda_1^{2*})$ otherwise, where λ_1^{2*} solves (B.4)

ii. Suppose $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 \geq 0$, i.e. we are at a situation as in top right chart of Figure 3. Following a symmetric argument, the characterization of this case follows:

Proposition 14. *Suppose $(\tilde{\lambda}_1^1 = 0, \tilde{\lambda}_1^2)$ with $\tilde{\lambda}_1^2 \leq 1$ and $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 \geq 0$. The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ if $\lambda_1^{2*} < \tilde{\lambda}_1^2$ where λ_1^{2*} solves (B.5).

ii. high-type partial PAM with $(1, \lambda_1^{2*})$ otherwise, where λ_1^{2*} solves (B.4).

Case 2. Suppose $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ with $\tilde{\lambda}_1^1 > 0$; i.e. we are at a situation as in the middle two chart of Figure 3. Now, we consider the sub-cases of $\hat{\lambda}_1^1$ as in the analyses above

i. Suppose $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$, i.e. we are at a situation as in the middle-left chart of Figure 3—which requires the slope to be less than -1. The following characterizes equilibrium sorting in this case:

Proposition 15. *Suppose $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ with $\tilde{\lambda}_1^1 > 0$ and $(\hat{\lambda}_1^1, \hat{\lambda}_1^2 = 0)$ with $\hat{\lambda}_1^1 \leq 1$. The equilibrium exhibits*

i. high-type partial NAM with $(0, \lambda_1^{2*})$ if $\lambda_1^{2*} < \tilde{\lambda}_1^2$ where λ_1^{2*} solves (B.5).

ii. *Suppose* $\lambda_1^{2*} \geq \tilde{\lambda}_1^2$. *Then the equilibrium exhibits*

a. *perfect NAM if* $\frac{\partial W(\lambda_1^1,\lambda_1^2)}{\partial \lambda_1^1}|_{\{\lambda_1^1=0,\lambda_1^2=1\}} \leq 0$

b. *low-type partial NAM with* $(\lambda_1^{1*}, 1)$, *if* $\lambda_1^{1*} \leq \tilde{\lambda}_1^1$ *where* λ_1^{1*} *solves* (B.7)

c. *Suppose* $\lambda_1^{1*} > \tilde{\lambda}_1^2$ *where* λ_1^{1*} *solves* (B.7). *Then the equilibrium exhibits*

[1.] *low-type partial PAM with* $(\lambda_1^{1*}, 0)$, *if* $\lambda_1^{1*} < 1$ *where* λ_1^{1*} *solves* (B.2)

[2.] *perfect PAM if* (B.3) *holds, and*

[3.] *high-type partial PAM with* $(1, \lambda_1^{2*})$, *otherwise, where* λ_1^{2*} *solves* (B.4)

ii. Suppose $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ with $\hat{\lambda}_1^2 \geq 0$, i.e., we are at a situation as in middle-right chart of Figure 3. Following a symmetric argument, the characterization of this case follows:

Proposition 16. *Suppose* $(\tilde{\lambda}_1^1, \tilde{\lambda}_1^2 = 1)$ *with* $\tilde{\lambda}_1^1 > 0$ *and* $(\hat{\lambda}_1^1 = 1, \hat{\lambda}_1^2)$ *with* $\hat{\lambda}_1^2 > 0$. *The equilibrium exhibits*

i. *high-type partial NAM with* $(0, \lambda_1^{2*})$ *if* $\lambda_1^{2*} < 1$ *where* λ_1^{2*} *solves* (B.5).

ii. *perfect NAM if* (B.6) *holds, and*

iii. *low-type partial NAM with* $(\lambda_1^{1*}, 1)$ *where* λ_1^{1*} *solves* (B.7), *otherwise.*