Online Supplement to "Designing the Menu of Licenses for Foster Care"

Altınok, Ahmet [*] MacDonald, Diana[†]

April 2026

Abstract

This document contains proofs and examples omitted from the main text.

A Appendix: Omitted Examples

In this section, we provide a more in-depth discussion of Examples 1 and 2. Additionally, we present environments that satisfy Corollaries 1 and 2, along with an explanation of how they fulfill the conditions for PAM.

A.1 Complete Information

Example A.1. (***Detailed explanation of Example 1***)**.** Figure A.1 illustrates an environment where super-modularity in the surplus function $S(x, y)$ is not a sufficient condition for PAM. We assume that the share of low-needs children is $f(x_1) = 0.8$, the functional form of the meeting technology is $\pi^p(\theta) = \frac{1}{1+\theta}$, and $S(x, y)$ is a super-modular function with values $S(x_2, y_2) = 191$, $S(x_1, y_2) = 201$, $S(x_2, y_1) = 40$ and $S(x_1, y_1) = 51$. Here, the condition over primitives presented in Corollary 1(i) is violated: $1 = \frac{S(x_2,y_2)-S(x_2,y_1)}{S(x_1,y_2)-S(x_1,y_1)} \not\geq \frac{1}{\pi^p(1/f(x_2))} = 5.99$

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

(a) Randomization Device (b) Market Tightness (c) Expected Transfers

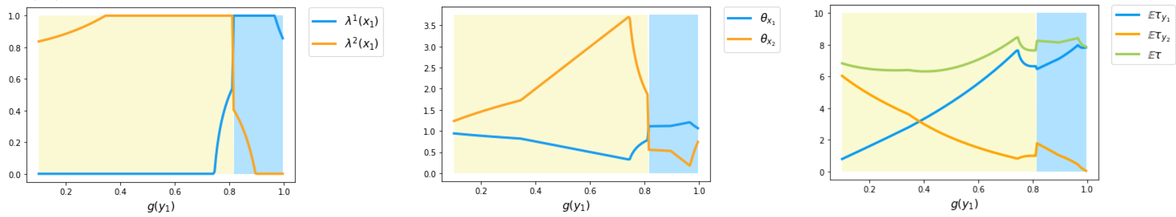


Figure A.1: Monotone Sorting Fails

Panel A.1a presents the optimal probability with which parents holding licenses 1 (blue line) and 2 (orange line) are allocated into submarket x_1. The y-axis corresponds to these probabilities while the x-axis presents possible values for the share of low-ability parents, $g(y_1)$. In Panel A.1b, we plot the optimal market tightness for submarket x_1 (blue line) and x_2 (orange line) as a function of the share of low-ability parents, $g(y_1)$. Here, the y-axis corresponds to possible values for the market tightness. In Panel A.1c, we plot the optimal expected transfers received by all y_1-parents (blue line) and all y_2-parents (orange line) as a function of the share of low-ability parents, $g(y_1)$. In addition, we also include the optimal total expected transfers (green line), or equivalently, the optimal total cost incurred by the child welfare agency to implement the optimal sorting. Lastly, in every graph, the blue- and golden- shaded regions correspond to PAM and NAM, respectively.

As Panel A.1a illustrates, for small enough values of $g(y_1)$, the equilibrium sorting exhibits NAM, even when the surplus function is super-modular. Thus, supermodularity is not a sufficient condition for PAM to hold in equilibrium. For the same interval of $g(y_1)$, Panel A.1b shows that the equilibrium market tightness is greater in submarket x_2 than x_1; thus, parents are more likely to meet a child in submarket x_1. This induces the designer to allocate y_2-parents in submarket x_1, resulting in NAM. As $g(y_1)$ increases, the equilibrium market tightness becomes larger in submarket x_1 than in x_2, and thus the equilibrium sorting reverses to PAM. Lastly, we can see from Panel A.1c that the total expected cost of imposing NAM increases as the share of low-ability parents increases.[1] This is intuitive, since low-ability parents incur a greater cost for providing care than high-ability parents. Therefore, the designer must pay greater

[1] Though the cost structure is not necessary for the analysis of equilibrium allocations, it directly determines the equilibrium total transfers. Panel A.1c uses the following cost function: $c(x_1, y_1) = 15, c(x_1, y_2) = 1, c(x_2, y_1) = 20, c(x_2, y_2) = 15$.

transfers to low-ability parents to satisfy the [PC]. □

Example A.2. (***Positive Assortative Matching Holds***)**.** Figure A.2 considers an environment that satisfies the condition presented in Corollary 1(i) to ensure that PAM will arise in equilibrium. In this environment, we assume that the share of low-needs children is equal to 0.8, and $S(x, y)$ is a super-modular function with revised values $S(x_2, y_2) = 100$, $S(x_1, y_2) = 201$, $S(x_2, y_1) = 30$ and $S(x_1, y_1) = 191$. This set of primitives satisfies the following: $7 = \frac{S(x_2,y_2)-S(x_2,y_1)}{S(x_1,y_2)-S(x_1,y_1)} \geq \frac{1}{\pi^p(1/f(x_2))} = 5.99$

(a) Randomization Device (b) Market Tightness (c) Expected Transfers

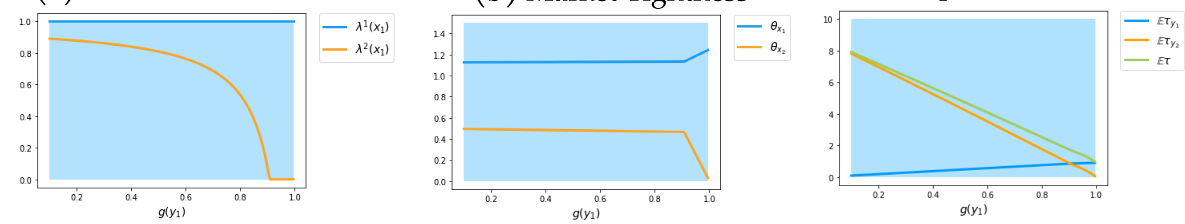


Figure A.2: Monotone Sorting Holds

As Panel A.2a illustrates, the equilibrium sorting exhibits PAM for any value of $g(y_1)$. Moreover, for sufficiently high values of $g(y_1)$, there is a perfect segregation of the market such that all type-y_i parents are allocated into submarket x_i.

Panel A.2b shows that the market tightness in both submarkets remains flat for a fair range of values of $g(y_1)$, even though the share of y_2-parents being allocated into submarket x_1 decreases. This is due to two effects compensating: **(i)** $\theta_1^\star$ increases as $g(y_1)$ increases, and **(ii)** $\theta_1^\star$ decreases as $\lambda^{2\star}(x_1)$ decreases. Similarly, for $\theta_2^\star$. In addition, note that the market tightness is larger in submarket x_1 than in submarket x_2, resulting in parents being more likely to meet a child in submarket x_2. This is in line with the intuition that the designer would like to allocate more profitable parents into thicker submarkets.

Lastly, in Panel A.2c, the total expected cost of implementing PAM is decreasing in $g(y_1)$, unlike the intuition presented in the previous example.[2] Here, the total expected transfers received by y_1-parents (blue line) are increasing in $g(y_1)$, but not enough to compensate for the decrease of the total expected transfers received by y_2-parents (orange line). □

[2] The cost structure here is as follows: $c(x_1, y_1) = 2, c(x_1, y_2) = 1, c(x_2, y_1) = 20, c(x_2, y_3) = 15$.

A.2 Private Information

Example A.3. (***Detailed explanation of Example* 2**). Figure A.3 illustrates the environment in Example A.1, where super-modularity in the surplus function is not a sufficient condition for PAM.[3] In all panels, the solid lines represent the equilibrium objects under the complete information, while the dashed lines correspond to the private information. The cost function is super-modular with the following values: $c(x_2, y_2) = 15$, $c(x_1, y_2) = 1$, $c(x_2, y_1) = 20$ and $c(x_1, y_1) = 15$. Notice, it guarantees the existence of a separating menu of licenses under NAM, whereas any equilibrium exhibiting PAM does not screen parents.

(**a**) Randomization Device (**b**) Market Tightness (**c**) Expected Transfers

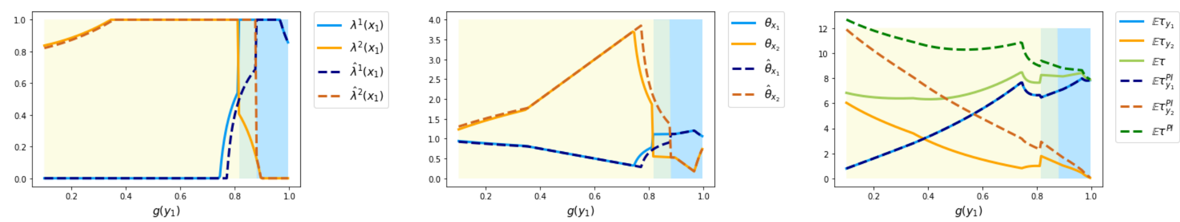


Figure A.3: Monotone Sorting Fails under Private Information

In Panel A.3a, one can observe that the optimal randomization devices, $\lambda^1(x_1)$ and $\lambda^2(x_2)$, are very similar for the complete and private information cases.[4] As a result, Panel A.3b represents that the optimal market tightness coincides at a fairly large interval for $g(y_1)$. However, when $g(y_1)$ is approximately in $(0.8, 0.9)$, the equilibrium sorting pattern is PAM under complete information, whereas it is NAM with private information. To see the intuition, consider an equilibrium menu of licenses that implements perfect sorting under complete information, such that $\tau^k(x) = c(x, y_k)$. If the menu implements NAM, then type-y_2 parent mimics y_1 and matches with child x_2 instead of x_1, and if it implements PAM, type-y_2 mimics y_1 and matches with x_1 instead of x_2. The former misreport allows parent y_2 to (ex-post) gain as much as $\tau^1(x_2) - c(x_2, y_2) = 5$ whereas the latter does $\tau^1(x_1) - c(x_1, y_2) = 14$. That is, y_2 has *stronger* incentives to misreport if the equilibrium sorting is PAM than when it is NAM. Thus, it is cheaper for the

[3] Recall that the primitives used in Example A.1 are as follows: $f(x_1) = 0.8$, $\pi^p(\theta) = \frac{1}{1+\theta}$, $S(x_2, y_2) = 191$, $S(x_1, y_2) = 201$, $S(x_2, y_1) = 40$ and $S(x_1, y_1) = 51$.

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

designer to switch the equilibrium sorting from PAM to NAM for the (roughly) specified region of $g(y_1)$. Notice, this intuition is in line with the counterpart of Corollary 2(i).

Lastly, in Panel A.3c we observe that the total expected transfers received by low-ability parents (blue lines) coincides under complete and private information. The reason is that, for low-ability parents, transfers are pint-down by the [PC] regardless of the information friction. For high-ability parents (orange lines), total expected transfers are greater under private than complete information. This is intuitive, since the designer must pay informational rents to incentivize high-ability parents to reveal their type truthfully when informational fictions are introduced. □

Example A.4. (***Positive Assortative Matching Holds***)**.** Figure A.4 illustrates the equilibrium objects of the environment in Example A.2 that satisfies the additional conditions presented in Corollary 3(i).[5] We assume $c(x, y)$ is a strong sub-modular function with values $c(x_2, y_2) = 13$, $c(x_1, y_2) = 1$, $c(x_2, y_1) = 20$ and $c(x_1, y_1) = 2$.

One can easily verify that the cost function guarantees the existence of a separating menu of licenses under PAM. In this case, the conditions over primitives presented in Corollary 3(i) are satisfied:

$$\frac{S(x_2, y_2) - S(x_2, y_1)}{S(x_1, y_2) - S(x_1, y_1)} = 7 = \frac{c(x_2, y_1) - c(x_2, y_2)}{c(x_1, y_1) - c(x_1, y_2)} \geq \frac{1}{\pi^p\left(1/f(x_2)\right)} = 5.99$$

As Panel A.4a illustrates, the equilibrium sorting exhibits PAM for any value of $g(y_1)$. Thus, PAM is robust to informational frictions, unlike Example A.3 where we observe PAM and NAM. In Panel A.4b, we observe that the market tightness in both submarkets remains flat for a fair range of values of $g(y_1)$. Lastly, in Panel A.4c we observe that the total cost of imposing PAM decreases with $g(y_1)$. Note that, in this example, the equilibrium allocations are almost identical under complete and private information.

[5] Recall that the primitives used in Example A.2 are as follows: $f(x_1) = 0.8$, $\pi^p(\theta) = \frac{1}{1+\theta}$, $S(x_2, y_2) = 100$, $S(x_1, y_2) = 201$, $S(x_2, y_1) = 30$ and $S(x_1, y_1) = 191$.

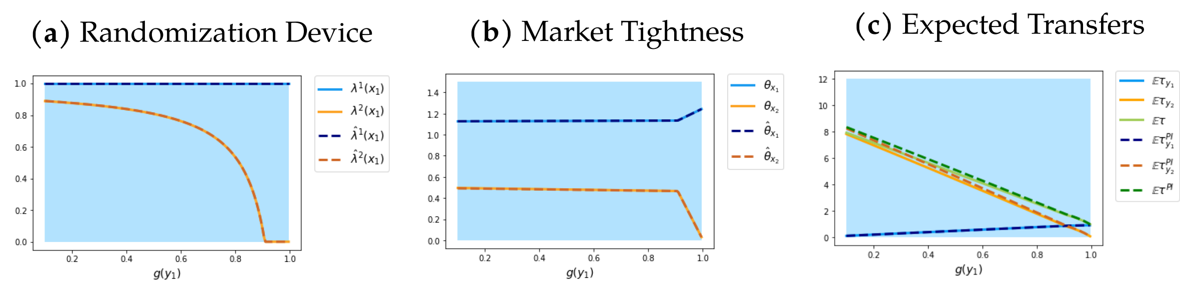


Figure A.4: Monotone Sorting Holds under Private Information

□

B Appendix: Assortative Matching in Equilibrium

B.1 Complete Information

Under the light of our results, we can characterize the equilibrium sorting patterns. We will start by providing some auxiliary lemmas.

Lemma B.1. *The rate of change in Welfare $W\big(\lambda^1(x_1), \lambda^2(x_1)\big)$ monotonically decreases in $\lambda^k(x_1)$ for each $k = 1, 2$.*

Proof. Recall the total welfare:

$$W\big(\lambda^1(x_1), \lambda^2(x_1)\big) = \pi^p(\theta_1) \cdot \underbrace{\Big[g(y_1)\, \lambda^1(x_1)\, S(x_1, y_1) + \big(1 - g(y_1)\big)\, \lambda^2(x_1)\, S(x_1, y_2)\Big]}_{\mathbb{E}U_1}$$

$$+ \pi^p(\theta_2) \cdot \underbrace{\Big[g(y_1)\, \big(1 - \lambda^1(x_1)\big)\, S(x_2, y_1) + \big(1 - g(y_1)\big)\, \big(1 - \lambda^2(x_1)\big)\, S(x_2, y_2)\Big]}_{\mathbb{E}U_2}$$

where

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

Fix $\lambda^{-k}(x_1)$. Increasing $\lambda^k(x_1)$ by a small amount $\varepsilon > 0$, increases $\mathbb{E}U_1$ and θ_1 linearly, and decreases $\mathbb{E}U_2$ and θ_2 linearly. Recall that, $\pi^p(\cdot)$ is a decreasing and convex function, thus the rate of increase through $\pi^p(\theta_1) \cdot \mathbb{E}U_1$ decreases, while the rate of decrease through $\pi^p(\theta_2) \cdot \mathbb{E}U_2$ increases in $\lambda^k(x_1)$, for any $k = 1, 2$. □

Lemma B.1 is useful since it implies that $\frac{\partial W\left(\lambda^1(x_1),\lambda^2(x_1)\right)}{\partial \lambda^k(x_1)}$ is monotonically decreasing. Thus, if it is zero at some $\lambda^{k\prime}(x_1)$, then it is negative at any $\lambda^k(x_1)$ if and only if $\lambda^k(x_1) > \lambda^{k\prime}(x_1)$ for any $\lambda^{-k}(x_1)$. Note that the same analysis applies to any pair $(\lambda^1(x_1), \lambda^2(x_1))$ that yields the same market tightness. Now, another useful lemma follows:

Lemma B.2. *Fix* $(\hat{\lambda}^1(x_1), \hat{\lambda}^2(x_1))$. *For any* $(\tilde{\lambda}^1(x_1), \tilde{\lambda}^2(x_1))$ *such that* $\theta_1(\hat{\lambda}^1(x_1), \hat{\lambda}^2(x_1)) = \theta_1(\tilde{\lambda}^1(x_1), \tilde{\lambda}^2(x_1))$ *and* $\hat{\lambda}^1(x_1) \geq \tilde{\lambda}^1(x_1)$, *the following holds:*

$$\frac{\partial W(\lambda^1(x_1), \lambda^2(x_1))}{\partial \lambda^k(x_1)}|_{(\hat{\lambda}^1(x_1),\hat{\lambda}^2(x_1))} \leq \frac{\partial W(\lambda^1(x_1), \lambda^2(x_1))}{\partial \lambda^k(x_1)}|_{(\tilde{\lambda}^1(x_1),\tilde{\lambda}^2(x_1))}$$

Proof. Taking partial derivatives on welfare yields the followings:

$$\frac{\partial W(\lambda^1(x_1), \lambda^2(x_1))}{\partial \lambda^1(x_1)} = g(y_1) \cdot V(\lambda^1(x_1), \lambda^2(x_1)) + g(y_1) \cdot \left[\pi^p(\theta_1)S(x_1, y_1) - \pi^p(\theta_2)S(x_2, y_1)\right]$$

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

with

$$V(\lambda^1(x_1), \lambda^2(x_1)) = \frac{\pi^{p\prime}(\theta_1)}{f(x_1)} \cdot \mathbb{E}U_1(\lambda^1(x_1), \lambda^2(x_1)) - \frac{\pi^{p\prime}(\theta_2)}{1 - f(x_1)} \cdot \mathbb{E}U_2(\lambda^1(x_1), \lambda^2(x_1))$$

where $\mathbb{E}U_1(\lambda^1(x_1), \lambda^2(x_1))$ and $\mathbb{E}U_2(\lambda^1(x_1), \lambda^2(x_1))$ are defined as in Lemma B.1. It is easy to verify that $V(\lambda^1(x_1), \lambda^2(x_1))$ decreases as we move down on the market tightness θ_1, that is, as we increase $\lambda^1(x_1)$ while decreasing $\lambda^2(x_1)$. This implies that the rate of change with respect to $\lambda^1(x_1)$ decreases as one moves down on the same market tightness, which finishes the proof. □

Now, by using Lemmas B.1 and B.2, we can characterize the equilibrium allocation of parents across submarket step by step. We establish the equilibrium allocation of parents when the sufficient conditions of Corollary 1 hold. Then we leave the analysis for the case where the sufficient conditions are violated to the readers.

Proposition B.1 (**Positive Assortative Matching (PAM)**). *Suppose* $\frac{S(x_2,y_2)-S(x_2,y_1)}{S(x_1,y_2)-S(x_1,y_1)} \geq \frac{1}{\pi^p\left(\frac{1}{f(x_2)}\right)}$ *holds. The equilibrium sorting exhibits:*

i. low-type PAM with $\lambda^{1\star}(x_1) \in (0,1)$ *and* $\lambda^{2\star}(x_1) = 0$ *if*

$$\frac{\partial W\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial \lambda^1(x_1)}|_{\{\lambda^2(x_1)=0\}} = 0 \text{ for some } \lambda^{1\star}(x_1) \in (0,1) \tag{B.1}$$

ii. perfect PAM with $\lambda^{1\star}(x_1) = 1$ *and* $\lambda^{2\star}(x_1) = 0$ *if*

$$\frac{\partial W\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial \lambda^1(x_1)}|_{\{\lambda^2(x_1)=1,\lambda^2(x_1)=0\}} \geq 0 \geq \frac{\partial W\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial \lambda^2(x_1)}|_{\{\lambda^1(x_1)=1,\lambda^2(x_1)=0\}} \tag{B.2}$$

iii. high-type PAM with $\lambda^{1\star}(x_1) = 1$ *and* $\lambda^{2\star}(x_1) \in (0,1)$ *if*

$$\frac{\partial W\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial \lambda^2(x_1)}|_{\{\lambda^1(x_1)=1\}} = 0 \text{ for some } \lambda^{2\star}(x_1) \in (0,1) \tag{B.3}$$

Proof. By assumption, $\frac{S(x_2,y_2)-S(x_2,y_1)}{S(x_1,y_2)-S(x_1,y_1)} \geq \frac{1}{\pi^p\left(\frac{1}{f(x_2)}\right)}$, which implies that $Z^{CI}(\theta_1) \geq 0$ for any θ_1. Therefore, starting from an initial allocation $\lambda^1(x_1) = 0$ and $\lambda^2(x_1) = 0$, the designer first allocates y_1-parents into submarket x_1 until either parents are exhausted or it is not profitable anymore. Accordingly, perfect PAM and high-type PAM follows. □

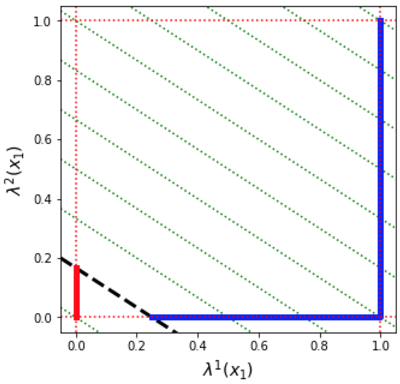
Proposition B.1 characterizes the equilibrium sorting patterns when the conditions specified in Corollary 1(i) hold. One can easily characterize the equilibrium distribution of parents across submarkets for NAM and extend it to the case where the sufficient conditions are violated, with a parallel argument. As is represented by the dashed-lines in Figure B.1, suppose that $Z^{CI}(\bar{\theta}_1) = 0$ for some $\bar{\theta}_1 \in (0, 1/f(x_1))$. Then, for NAM (red-lines in Figure B.1), there exists either, **(i)** $\tilde{\lambda}^1(x_1) = 0$ and $\tilde{\lambda}^2(x_1) \leq 1$ or **(ii)** $\tilde{\lambda}^1(x_1) > 0$ and $\tilde{\lambda}^2(x_1) = 1$, with $\bar{\theta}_1 = \frac{g(y_1)\tilde{\lambda}^1(x_1)+(1-g(y_1))\tilde{\lambda}^2(x_1)}{f(x_1)}$. Similarly, for PAM (blue-lines in Figure B.1), there exists either, **(i)** $\hat{\lambda}^1(x_1) \leq 1$ and $\hat{\lambda}^2(x_1) = 0$ or **(ii)** $\hat{\lambda}^1(x_1) = 1$ and $\hat{\lambda}^2(x_1) \geq 0$, with $\bar{\theta}_1 = \frac{g(y_1)\hat{\lambda}^1(x_1)+(1-g(y_1))\hat{\lambda}^2(x_1)}{f(x_1)}$. In what follows, we study each possible case illustrated in Figure B.1.

B.2 Private Information

Lemmas B.1 and B.2 carry over to the case of private information:

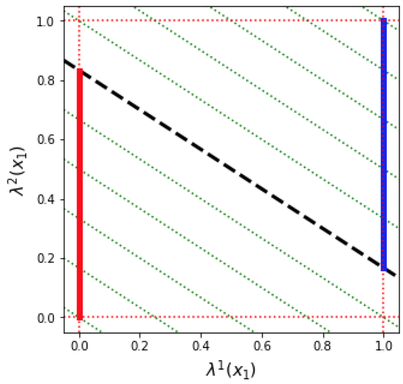
Figure B.1: Possible Cases given $Z^{CI}(\overline{\theta}_1)$

(a) Case 1A



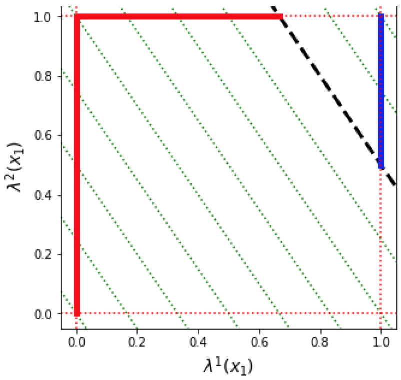
$\tilde{\lambda}^1(x_1) = 0$ and $\tilde{\lambda}^2(x_1) \leq 1$
$\hat{\lambda}^1(x_1) \leq 1$ and $\hat{\lambda}^2(x_1) = 0$

(b) Case 1B



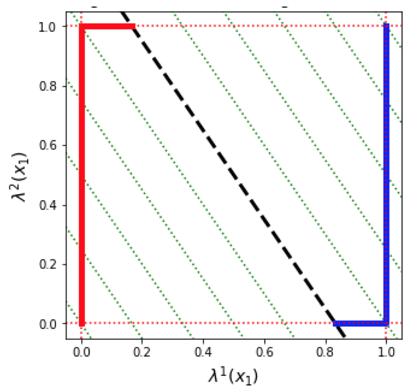
$\tilde{\lambda}^1(x_1) = 0$ and $\tilde{\lambda}^2(x_1) \leq 1$
$\hat{\lambda}^1(x_1) = 1$ and $\hat{\lambda}^2(x_1) \geq 0$

(c) Case 2A



$\tilde{\lambda}^1(x_1) \geq 0$ and $\tilde{\lambda}^2(x_1) = 1$
$\hat{\lambda}^1(x_1) = 1$ and $\hat{\lambda}^2(x_1) \geq 0$

(d) Case 2B



$\tilde{\lambda}^1(x_1) \geq 0$ and $\tilde{\lambda}^2(x_1) = 1$
$\hat{\lambda}^1(x_1) \leq 1$ and $\hat{\lambda}^2(x_1) = 0$

Lemma B.3. *The rate of change in Welfare $\hat{W}\big(\lambda^1(x_1), \lambda^2(x_1)\big)$ monotonically decreases in $\lambda^k(x_1)$ for each $k = 1, 2$.*

Proof. Recall the welfare of children:

$$\hat{W}\big(\lambda^1(x_1), \lambda^2(x_1)\big) = \pi^p(\theta_1) \cdot \underbrace{\Big\{ g(y_1)\, \lambda^1(x_1)\, S(x_1, y_1) + \big(1 - g(y_1)\big) \Big[\lambda^2(x_1)\, S(x_1, y_2) - \lambda^1(x_1) \Big(c(x_1, y_1) - c(x_1, y_2) \Big) \Big] \Big\}}_{\mathbb{E}\hat{U}_1}$$

$$+ \pi^p(\theta_2) \cdot \underbrace{\Big\{ g(y_1)\, \big(1 - \lambda^1(x_1)\big)\, S(x_2, y_1) + \big(1 - g(y_1)\big) \Big[\big(1 - \lambda^2(x_1)\big)\, S(x_2, y_2) - \big(1 - \lambda^1(x_1)\big) \Big(c(x_2, y_1) - c(x_2, y_2) \Big) \Big] \Big\}}_{\mathbb{E}\hat{U}_2}$$

where

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

Fix $\lambda^{-k}(x_1)$. Increasing $\lambda^k(x_1)$ by a small amount $\varepsilon > 0$, increases $\mathbb{E}\hat{U}_1$ and θ_1, and decreases $\mathbb{E}\hat{U}_2$ and θ_2 linearly. Since $\pi^p(\cdot)$ is a decreasing and convex function, the rate of increase through $\pi^p(\theta_1)\cdot\mathbb{E}\hat{U}_1$ decreases, while the rate of decrease through $\pi^p(\theta_2)\cdot\mathbb{E}\hat{U}_2$ increases in $\lambda^k(x_1)$, for any $k = 1, 2$. □

Lemma B.3 implies that $\frac{\partial \hat{W}\left(\lambda^1(x_1),\lambda^2(x_1)\right)}{\partial \lambda^k(x_1)}$ is monotonically decreasing. Now, another useful lemma follows:

Lemma B.4. *Fix $\big(\hat{\lambda}^1(x_1), \hat{\lambda}^2(x_1)\big)$. For any $\big(\tilde{\lambda}^1(x_1), \tilde{\lambda}^2(x_1)\big)$ such that $\theta_1\big(\hat{\lambda}^1(x_1), \hat{\lambda}^2(x_1)\big) = \theta_1\big(\tilde{\lambda}^1(x_1), \tilde{\lambda}^2(x_1)\big)$ and $\hat{\lambda}^1(x_1) \geq \tilde{\lambda}^1(x_1)$, the following holds:*

$$\frac{\partial \hat{W}\big(\lambda^1(x_1), \lambda^2(x_1)\big)}{\partial \lambda^k(x_1)}\Big|_{\left(\hat{\lambda}^1(x_1),\hat{\lambda}^2(x_1)\right)} \leq \frac{\partial \hat{W}\big(\lambda^1(x_1), \lambda^2(x_1)\big)}{\partial \lambda^k(x_1)}\Big|_{\left(\tilde{\lambda}^1(x_1),\tilde{\lambda}^2(x_1)\right)}$$

Proof. Taking partial derivative on welfare under private information yields the followings:

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

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

with

$$V(\lambda^1(x_1), \lambda^2(x_1)) = \frac{\pi^{p\prime}(\theta_1)}{f(x_1)}\cdot\mathbb{E}\hat{U}_1(\lambda^1(x_1), \lambda^2(x_1)) - \frac{\pi^{p\prime}(\theta_2)}{1 - f(x_1)}\cdot\mathbb{E}\hat{U}_2(\lambda^1(x_1), \lambda^2(x_1))$$

where $\mathbb{E}\hat{U}_1(\lambda^1(x_1), \lambda^2(x_1))$ and $\mathbb{E}\hat{U}_2(\lambda^1(x_1), \lambda^2(x_1))$ are defined as in Lemma B.3. Notice, plugging $V(\lambda^1(x_1), \lambda^2(x_1))$ into $\frac{\partial \hat{W}\left(\lambda^1(x_1),\lambda^2(x_1)\right)}{\partial \lambda^1(x_1)}$ yields the following:

$$\frac{\partial \hat{W}\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial \lambda^1(x_1)} = \frac{g(y_1)}{\left(1-g(y_1)\right)} \cdot \frac{\partial \hat{W}\left(\lambda^1(x_1), \lambda^2(x_1)\right)}{\partial \lambda^2(x_1)} + g(y_1) \cdot Z^{PI}(\theta_1)$$

It is easy to see that $V(\lambda^1(x_1), \lambda^2(x_1))$ is the same as the complete information case, and thus, it decreases as we move down on the market tightness θ_1. In other words, as we increase $\lambda^1(x_1)$ while decreasing $\lambda^2(x_1)$, $V(\lambda^1(x_1), \lambda^2(x_1))$ decreases. This implies that the rate of change with respect to $\lambda^1(x_1)$ decreases as one moves down on the same market tightness, which finishes the proof. □

Now, by using Lemmas B.3 and B.4, we characterize the equilibrium allocation of parents across submarkets as in the complete information case.

Proposition B.2 (**Positive Assortative Matching(PAM)**). *Suppose that* $\frac{S(x_2,y_2)-S(x_2,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_2,y_1)-c(x_2,y_2)]}{S(x_1,y_2)-S(x_1,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_1,y_1)-c(x_1,y_2)]} \geq \frac{1}{\pi^p\left(\frac{1}{f(x_2)}\right)}$ *holds. The equilibrium exhibits:*

i. low-type PAM with $\lambda^{1\star}(x_1) \in (0,1)$ *and* $\lambda^{2\star}(x_1) = 0$ *if*

$$\frac{\partial \hat{W}(\lambda^1(x_1), \lambda^2(x_1))}{\partial \lambda^1(x_1)}|_{\{\lambda^2(x_1)=0\}} = 0 \text{ for some } \lambda^{1\star}(x_1) \in (0,1) \tag{B.4}$$

ii. perfect PAM with $\lambda^{1\star}(x_1) = 1$ *and* $\lambda^{2\star}(x_1) = 0$ *if*

$$\frac{\partial \hat{W}(\lambda^1(x_1), \lambda^2(x_1))}{\partial \lambda^1(x_1)}|_{\{\lambda^1(x_1)=1,\lambda^2(x_1)=0\}} \geq 0 \geq \frac{\partial \hat{W}(\lambda^1(x_1), \lambda^2(x_1))}{\partial \lambda^2(x_1)}|_{\{\lambda^1(x_1)=1,\lambda^2(x_1)=0\}} \tag{B.5}$$

iii. high-type PAM with $\lambda^{1\star}(x_1) = 1$ *and* $\lambda^{2\star}(x_1) \in (0,1)$ *if*

$$\frac{\partial \hat{W}(\lambda^1(x_1), \lambda^2(x_1))}{\partial \lambda^2(x_1)}|_{\{\lambda^1(x_1)=1\}} = 0 \text{ for some } \lambda^2(x_1)^* \in (0,1) \tag{B.6}$$

Proof. By assumption, $\frac{S(x_2,y_2)-S(x_2,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_2,y_1)-c(x_2,y_2)]}{S(x_1,y_2)-S(x_1,y_1)+\frac{g(y_2)}{g(y_1)}\cdot[c(x_1,y_1)-c(x_1,y_2)]} \geq \frac{1}{\pi^p\left(\frac{1}{f(x_2)}\right)}$ holds, which implies that $Z^{PI}(\theta_1) \geq 0$ for any θ_1. Therefore, starting from an initial allocation $\lambda^1(x_1) = 0$ and $\lambda^2(x_1) = 0$, the designer first allocated y_1-parents into submarket x_1 until either parents are exhausted or itis not profitable anymore. Accordingly, perfect PAM and high-type PAM follows. □

Proposition B.2 characterizes the equilibrium sorting patterns when the conditions specified in Corollary2(i) hold. We leave the analysis of the cases where the sufficient conditions are violated to the readers.