Supplemental Appendix
Endogenous Tracking: Sorting and Peer Effects

Aleksei Chernulich [a,*] Romain Gauriot [b,†] Daehong Min [c,‡]

A Model with Complete Information and Alternative Peer Effects Production Function

In this section, we study an alternative model of endogenous tracking. Our goal is to show that endogenous tracking can be rationalised without decreasing differences in costs, which is the main force supporting a sorting equilibrium in our model from the main text.

Compared to the model in the main text, the alternative model has different assumptions on (i) cost functions, (ii) the peer effect function, and (iii) the information structure. In the main text, we primarily hinge on the ability-dependent cost function—thus, decreasing differences in costs—to construct a sorting equilibrium. The alternative model has an ability-independent cost function, but we hinge on a different peer effect function to construct a sorting equilibrium. Lastly, we study a complete information game for clearer expositions.[1]

Consider n-players simultaneous-move game. Each player chooses one of two languages, $l \in \{G, S\}$. Each player is endowed with her own "ability" denoted by a_i, which is publicly known. Thus, the game is of complete information. Without loss of generality, assume that players are indexed from the highest to lowest ability in descending order so that a player with a higher index has a higher endowed ability: $a_n > a_{n-1} > \cdots > a_1$. We will refer to the graph of these ranked abilities as the "ability curve." Each language has an (opportunity) cost denoted by c_l: e.g., c_G represents the cost associated with learning the language G.

Assumption A.1. $c_G > c_S$ for all player i.

Note that the costs are independent of players' abilities. Thus, the cost difference between the two languages is the same for every player with any level of endowed ability. In the model from the main text, the cost difference decreases in each player's realised ability, a_i, which was essential to construct a non-trivial sorting equilibrium. However, as will

*Department of Economics, Durham University, *E-mail:* Aleksei.Chernulich@durham.ac.uk.

†Department of Economics, Deakin University, *E-mail:* Romain.Gauriot@deakin.edu.au.

‡Department of Economics, New York University Abu Dhabi and Korea Information Society Development Institute, *E-mail:* Daehong.Min@nyu.edu.

[1] One can study the same model in this section with incomplete information and provide conditions under which a sorting equilibrium can be constructed. However, the conditions are more complicated but do not provide more intuition beyond the complete information model in this section; they just obscure the intuition.

be clear soon, we do not need the decreasing differences in two costs in this alternative game due to a different form of the peer effect function.[2]

Denote by g_l the group of players who choose $l \in \{G, S\}$ as in the main text.

Assumption A.2. *Given any group composition, g_G and g_S, the peer effect is*

$$\mathcal{P}(g_l) = \frac{1}{|g_l|}\left(\sum_{j \in g_l} a_j\right) \quad \textit{for } l \in \{G, S\}.$$

Not as in the main text, the peer effect that player i experiences in g_l is the average of players' endowed abilities in g_l *including* her own ability. This different form of the peer effect function allows us to construct a sorting equilibrium without the ability-dependent cost functions. It is worth noting that we do not have the subscript i in the peer effect anymore. Since the peer effect player i experiences in g_l is simply the group average, every player in the same group experiences the same peer effect. More importantly, this entirely changes the discussion on a player's incentive to deviate to the other group. If player i deviates from one group to the other group, her ability can boost up or down the peer effect she would experience from her deviation, depending on whether her ability is above the average of the group she would deviate to. For example, given a group composition, suppose player i in g_G deviates to g_S. Her deviation changes not only the group composition but also the peer effect of each group, mainly due to Assumption A.2. Before her deviation, the peer effect in g_S is $\mathcal{P}(g_S) = \frac{1}{|g_S|}\sum_{j \in g_S} a_j$. After her deviation, the peer effect that player i experiences in g_S is not merely $\mathcal{P}(g_S)$ but $\mathcal{P}(g_S \cup \{i\}) = \frac{1}{|g_S \cup \{i\}|}\sum_{k \in g_S \cup \{i\}} a_k$. This contrasts with the discussion in the main text. In the main text, the peer effect that player i would experience when deviating to the other group is merely the average of abilities in that group, as her own ability is not included in the peer effect.

Each player in each group enjoys a positive peer effect and bears the cost of learning the chosen language. Each player's payoff is her net educational benefit, as in the main text. Hence, the payoff of player i in g_l,

$$u_i(l, g_l) = \mathcal{P}_i(g_l) - c_l.$$

Given any strategy profile, we refer to the player whose ability is the highest in g_l as player l for $l = G, S$.

Lemma A.1. *Given any group composition, if both players G and S do not want to deviate to the other group, it is a Nash equilibrium outcome.*

Lemma A.1 is intuitive. The player whose ability is the highest in each group contributes the most to the peer effect. In other words, the player whose ability is the highest in each group benefits the least from her peers. If the player who benefits the least from her peers does not have an incentive to deviate, the other players who benefit more than her would not deviate either.[3]

[2] If we have decreasing differences in costs, it is easier to construct a sorting outcome.

[3] Or, putting it differently, the player whose ability is the highest has the largest incentive to deviate to the other group because her ability is included in the other group's peer effect when deviating to that group and, thus, she is the player who would expect the highest peer effect from her deviation.

To prove Lemma A.1, consider each player's incentive to deviate from the group she belongs to. Player i in g_G does not want to deviate if and only if

$$\begin{aligned} & \mathcal{P}(g_G) - c_G \geq \mathcal{P}(g_S \cup \{i\}) - c_S \\ \iff \quad & \mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{i\}) \geq c_G - c_S. \end{aligned}$$

Note that $\mathcal{P}(g_S \cup \{i\})$ increases in player i's ability (or the player's index). Hence, we must have

$$\mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{i\}) \geq \mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{j\}) \quad \text{for } i < j.$$

Hence, if player G in g_G does not have an incentive to deviate, any player in g_G does not as well:

$$\mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{G\}) \geq \Delta c \implies \mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{i\}) \geq \Delta c \quad \text{for all } i \in g_G,$$

where $\Delta c = c_G - c_S$.

Now consider player j in g_S. Any player j in g_S does not have an incentive to deviate to g_G if and only if

$$\begin{aligned} & \mathcal{P}(g_S) - c_S \geq \mathcal{P}(g_G \cup \{j\}) - c_G \\ \iff \quad & \Delta c \geq \mathcal{P}(g_G \cup \{j\}) - \mathcal{P}(g_S). \end{aligned}$$

Again $\mathcal{P}(g_G \cup \{j\})$ increases in player j's ability. Similar to the previous case, we must have

$$\Delta c \geq \mathcal{P}(g_G \cup \{S\}) - \mathcal{P}(g_S) \implies \Delta c \geq \mathcal{P}(g_G \cup \{j\}) - \mathcal{P}(g_S) \text{ for all } j \in g_S.$$

That is, if player S in g_S does not have an incentive to deviate, any player in g_S does not as well.

It is worth noting that the difference in two peer effects—either $\mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{i\})$ for $i \in g_G$ or $\mathcal{P}(g_G \cup \{j\}) - \mathcal{P}(g_S)$ for $j \in g_S$—depends on each player's ability but the difference in two costs, Δc, is a constant for all players. This contrasts with the model in the main text: in the main text, the difference in two costs depends on the player's realised ability, but the difference in two peer effects is constant for all players. Thus, while we hinge on the ability-dependent cost functions to construct a sorting equilibrium in the main text, we use the fact that the difference in two peer effects depends on players' abilities to construct a sorting equilibrium in this alternative model.

Consider the following group composition in which the players are monotonically sorted:

$$g_S = \{a_1, \ldots, a_k\} \text{ and } g_G = \{a_{k+1}, \ldots, a_n\}.$$

We refer to such a group composition as a k-monotonic-sorting as the k-th player is pivotal. A k-monotonic-sorting is a Nash equilibrium outcome if and only if the player with the highest ability in each group does not have an incentive to deviate to the other group to which she does not belong by Lemma A.1. Thus, a given k-monotonic-sorting is a Nash equilibrium outcome if and only if both player n and player k have no incentive

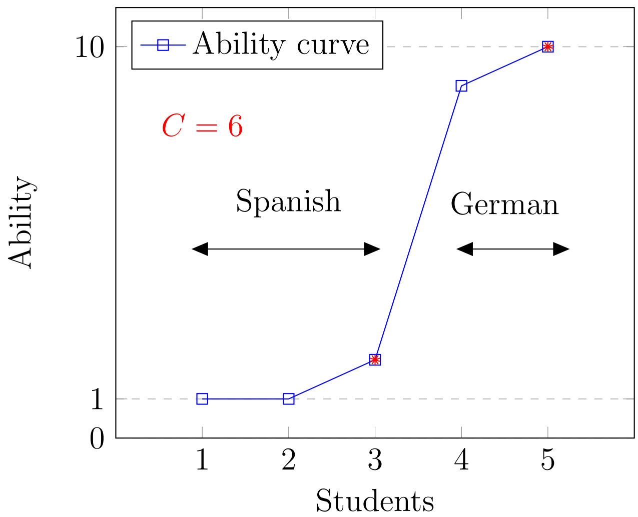

Figure 1: An example of the sorting equilibrium in a cohort of students with abilities $\{1, 1, 2, 9, 10\}$ and relative cost of choosing the German language is $\Delta c = 6$. Students with the highest ability in each group are marked with red–their incentives not to deviate are ensured by the relative cost supporting an equilibrium.

to deviate:

$$\mathcal{P}(g_G \cup k) - \mathcal{P}(g_S) \leq \Delta c \leq \mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{n\}).$$

Proposition A.1 summarises the discussion so far.

Proposition A.1. *A k-monotonic-sorting is a Nash equilibrium outcome if and only if* $\dfrac{a_k + a_{k+1} + \cdots + a_n}{n-k+1} - \dfrac{a_1 + \cdots + a_k}{k} \leq \Delta c \leq \dfrac{a_{k+1} + \cdots + a_n}{n-k} - \dfrac{a_1 + \cdots + a_k + a_n}{k+1}$.

Intuitively, for a k-monotonic-sorting to be supported as a Nash equilibrium outcome, it is sufficient (but not necessary) to have enough between-group heterogeneity and within-group homogeneity as shown in the following example.

Example. Consider the cohort of students with the following set of abilities: $\{1, 1, 2, 9, 10\}$. Further, assume that $c_G - c_S = \Delta c = 6$. With Proposition A.1, we can check if the following monotonic sorting is a Nash equilibrium outcome of the game:

$$g_S = \{1, 1, 2\} \quad \text{and} \quad g_G = \{9, 10\}.$$

The group composition given above is the 3-monotonic-sorting. Note that

$$\begin{aligned} \mathcal{P}(g_S \cup \{3\}) - \mathcal{P}(g_G) &= \frac{a_3 + a_4 + a_5}{3} - \frac{a_1 + a_2 + a_3}{3} \\ &= \frac{2 + 9 + 10}{3} - \frac{1 + 1 + 2}{3} \approx 5.7, \end{aligned}$$

and,

$$\begin{aligned} \mathcal{P}(g_G) - \mathcal{P}(g_S \cup \{5\}) &= \frac{a_4 + a_5}{2} - \frac{a_1 + a_2 + a_3 + a_5}{4} \\ &= \frac{9 + 10}{2} - \frac{1 + 1 + 2 + 10}{4} = 6. \end{aligned}$$

It is easy to check that the condition in Proposition A.1 holds, i.e.,

$$\frac{2+9+10}{3} - \frac{1+1+2}{3} \leq \Delta c \leq \frac{9+10}{2} - \frac{1+1+2+10}{4}$$

Hence, $G_g = \{9, 10\}$ and $G_s = \{1, 1, 2\}$ is a Nash equilibrium outcome outcome.

B Theoretical Observational Equivalence with a Signalling Model

This section provides a signalling model that is based on the assumptions used in the costly peer-seeking model from the main text. We keep the same assumptions on the model primitives such as $F(a)$ and $c_l(\lambda, a_i)$, but the players' utility functions take a different form and, in this case, do not include peer effects. We also introduce an employer who wants to hire the players and can use language class choices as informative signals.

There are n players. There is commonly known distribution, $F(a)$, over the set of possible abilities, $A = [\underline{a}, \bar{a}]$. Each player observes her own ability, a_i, and decides to choose G or S. Each action costs the player differently, as in the main text. That is, $\Delta c(\lambda, a_i) > 0$ and $\partial \Delta c(\lambda, a_i)/\partial a_i < 0$ for all $a_i \in A$.

Suppose that there is an employer who wants to hire an employee. The employer wants to maximise the profit that he can make from his hiring decision. Assume that the employer's profit is the difference between an employee's true ability and the wage paid to the employee:

$$u^E(a_i, w) = a_i - w.$$

Assuming that the potential employees are the players above, the game proceeds as follows. First, the players simultaneously decide their actions. The employer cannot observe the true abilities of potential employees. Instead, the employer can observe whether each potential employee chooses G or S and decide the wage for each employee. Thus, the employer's strategy is

$$s_E : \{G, S\} \rightarrow \mathbb{R}_+,$$

which boils down to a pair of real numbers, (w_G, w_S), i.e., a wage for each action. For simplicity, we assume that the employer does not have budget constraints and can hire all players.

Given the strategy of the employer, each player's payoff is

$$u^P(l) = w_l - c_l(\lambda, a_i) \quad \text{for} \quad l \in \{G, S\},$$

and each player's strategy is $s_i : A \rightarrow \{G, S\}$. Now, note that each player's payoff does not depend on the other players' strategies. Hence, we have n independent games, where the employer and each player i play each independent game.

Assume that the labour market is perfectly competitive, and, thus, the employer can only make zero profit. Then, given a player's strategy, s_i, the employer observes either G or S. Then, the employer updates his beliefs about the ability of the player and makes a wage offer that makes his profit zero at an equilibrium: i.e., $w_l = E[a_i | l, s_i]$ for $l \in \{G, S\}$.

Now, suppose that each player plays a threshold strategy as in the main text:

$$\hat{s}_i = \begin{cases} G & \text{if } a_i \geq \hat{a}, \\ S & \text{otherwise.} \end{cases}$$

Given $\hat{s}_i$, we have

$$E[a_i|G, \hat{s}_i] = E[a_i|a_i \geq \hat{a}], \quad \text{and} \quad E[a_i|S, \hat{s}_i] = E[a_i|a_i < \hat{a}],$$

which are the wages offered by the employer. Then, given a realised ability a_i, player i's payoff is

$$\begin{aligned} u^P(l; a_i) &= E[a_i|a_i \geq \hat{a}] - c_G(\lambda, a_i) \quad \text{for} \quad l = G, \\ &= E[a_i|a_i < \hat{a}] - c_S(\lambda, a_i) \quad \text{for} \quad l = S, \end{aligned}$$

which is equivalent to each player's (expected) payoff in the model from the main text. Note that here we have simply replaced the expected peer effects with the employer's wage offer. Thus, player i's payoff difference in this alternative setup, $\Delta u^P(a_i)$, is equal to player i's expected payoff difference when the other players play the same threshold strategies, $\hat{\boldsymbol{s}}_{-i}$, in the main text, $\Delta E[u_i(a_i|\hat{\boldsymbol{s}}_{-i})]$:

$$\begin{aligned} \Delta u^P(a_i) &= w_G - w_S - (c_G(\lambda, a_i) - c_S(\lambda, a_i)) = E[a_i|a_i \geq \hat{a}] - E[a_i|a_i < \hat{a}] - \Delta c(\lambda, a_i) \\ &= \Delta E[\mathcal{P}_i(\hat{a})] - \Delta c(\lambda, a_i) = \Delta E[u_i(a_i|\hat{\boldsymbol{s}}_{-i})]. \end{aligned}$$

Then, all the results in the main text are true in this alternative model. Proposition 1 provides the sufficient conditions for an equilibrium with s_i^* to exist, where $s_i^* = G$ for $a_i \geq a^*$ and $s_i^* = S$ for $a_i < a^*$ for $a^* \in (\underline{a}, \bar{a})$. Proposition 2 provides the comparative statics results.

Note that now we have n Perfect Bayesian Nash Equilibria for n independent games, where each player plays the threshold strategy with $\hat{a} = a^*$. Then, the aggregated outcome of these n signalling games is equivalent to the equilibrium we have in the language choice game with a peer-seeking motive.

C Histograms of genetic matching estimates

Genetic matching heavily relies on the genetic algorithm to find the weights that are used for matching. The algorithm mimics the evolutionary process and includes a random component. For that reason, the results of the genetic matching also include some randomness. We extended the original genetic matching approach and ran the estimation for the genetic matching a thousand times. Figures 2, 3, 4, and 5 show histograms of the 1000 estimates we get from these 1000 iterations.

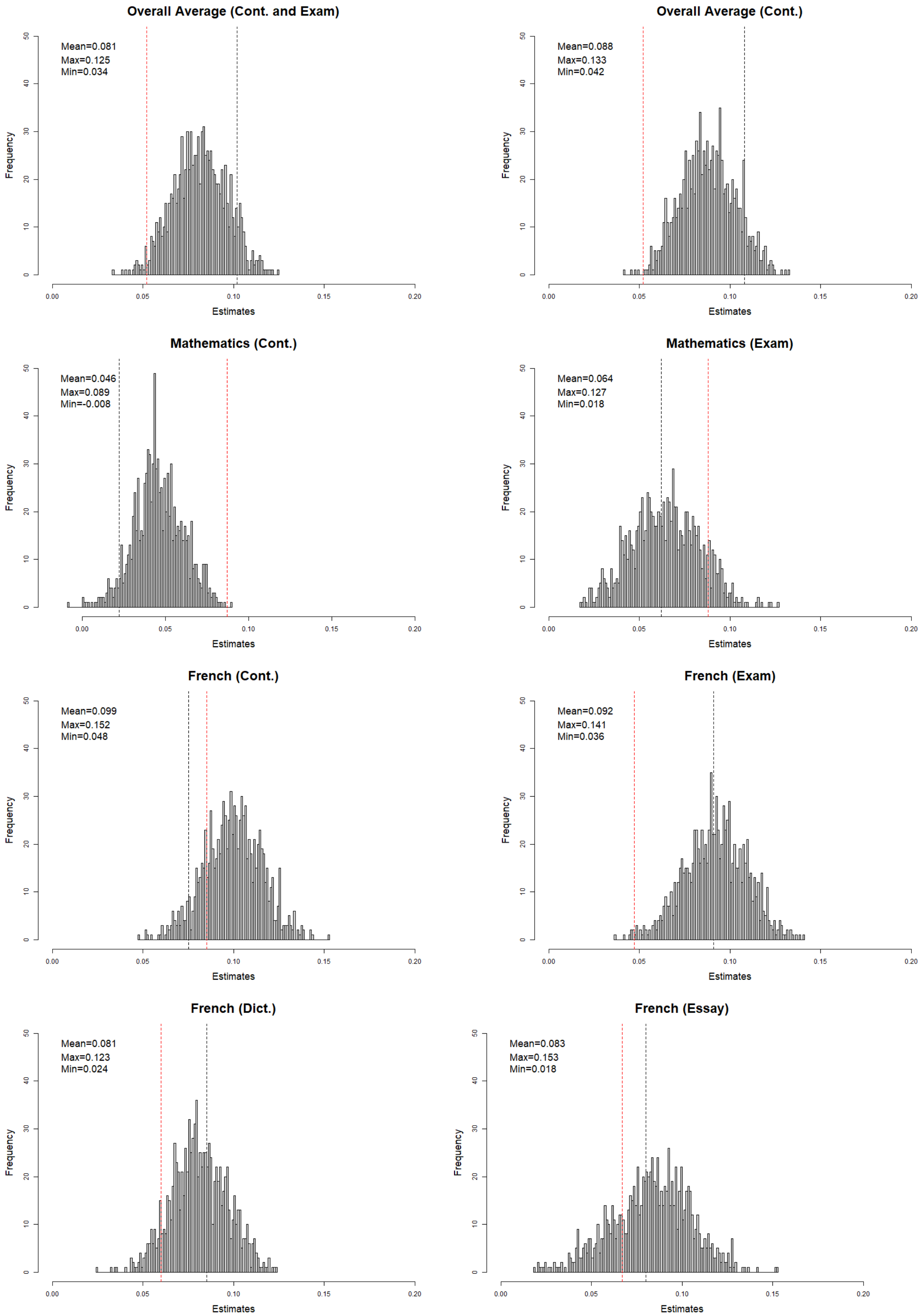

Figure 2: Histograms of 1000 realisations of the genetic matching estimates for the Difference in grade in Brevet des College exam in 2011 with a population size of 10,000 when matching on all covariates (fifth column of Table 8). The black line represents the estimate from genetic matching shown in Table 8, and the red line is the propensity score estimate shown in Table 14.

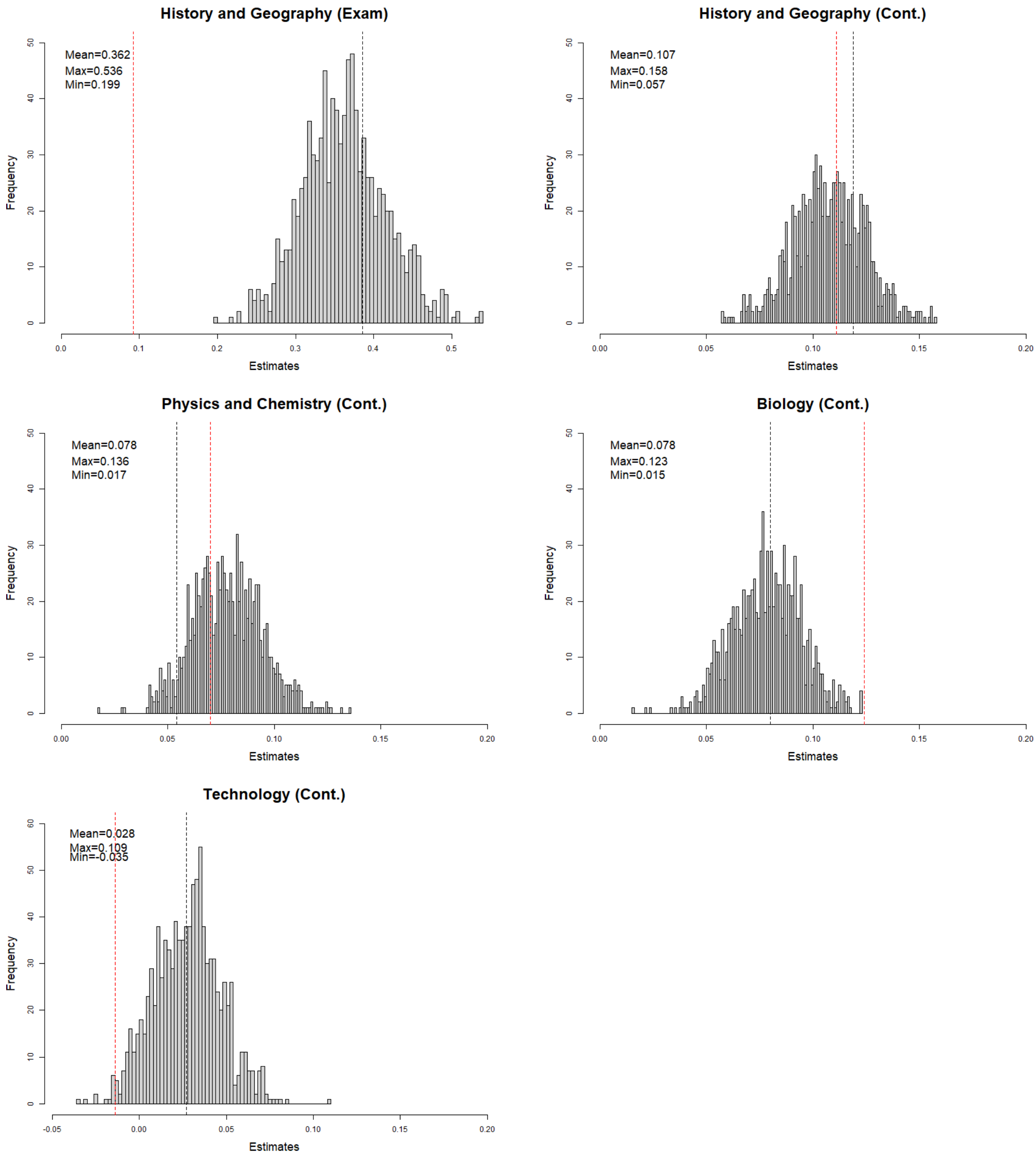

Figure 3: Histograms of 1000 realisations of the genetic matching estimates for the Difference in grade in Brevet des College exam in 2011 with a population size of 10,000 when matching on all covariates (fifth column of Table 9). The black line represents the estimate from genetic matching shown in Table 9, and the red line is the propensity score estimate shown in Table 15.

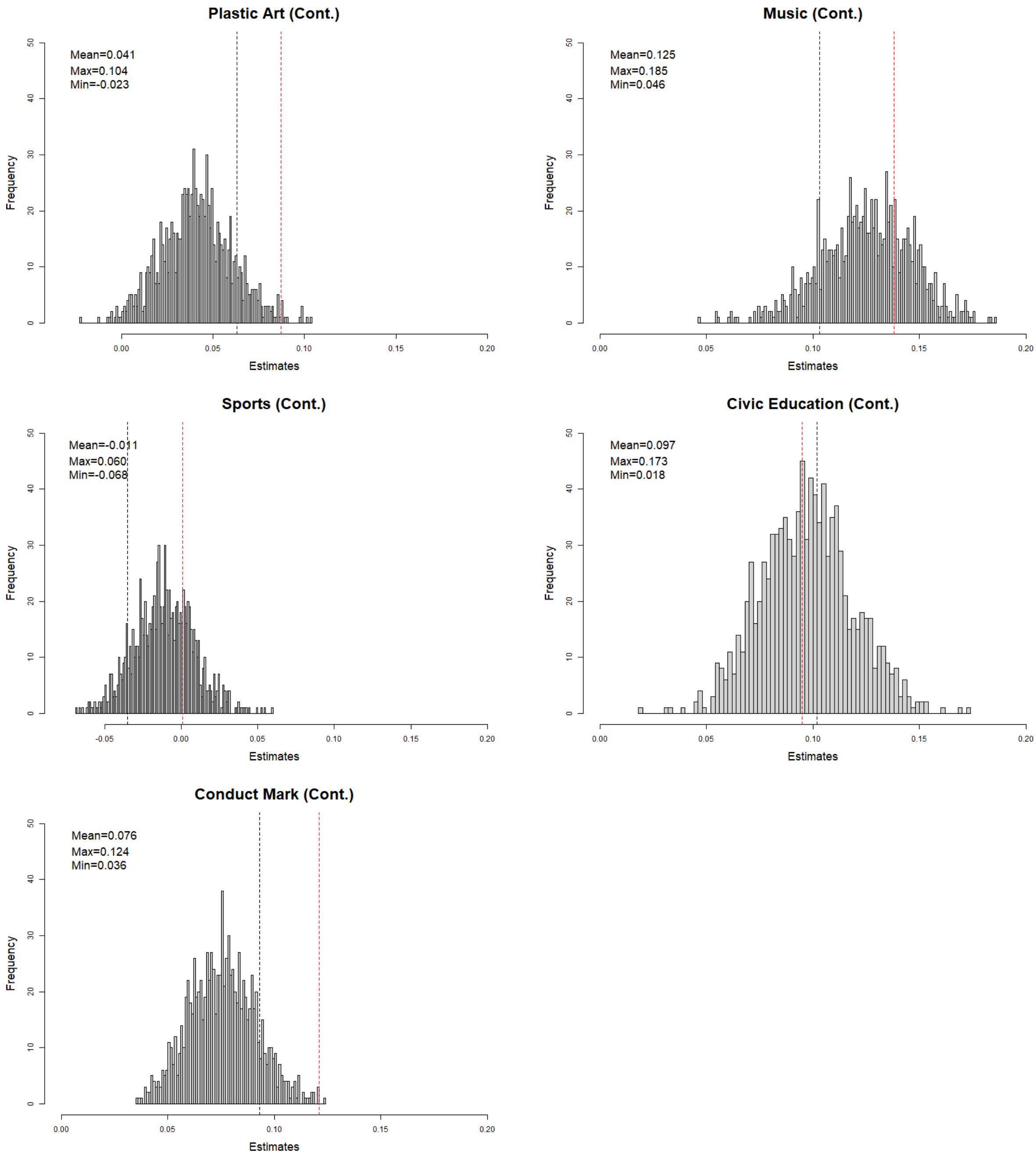

Figure 4: Histograms of 1000 realizations of the genetic matching estimates for the Difference in grade in Brevet des College exam in 2011 with a population size of 10,000 when matching on all covariates (fifth column of Table 10). The black line represents the estimate from genetic matching shown in Table 10, and the red line is the propensity score estimate shown in Table 16.

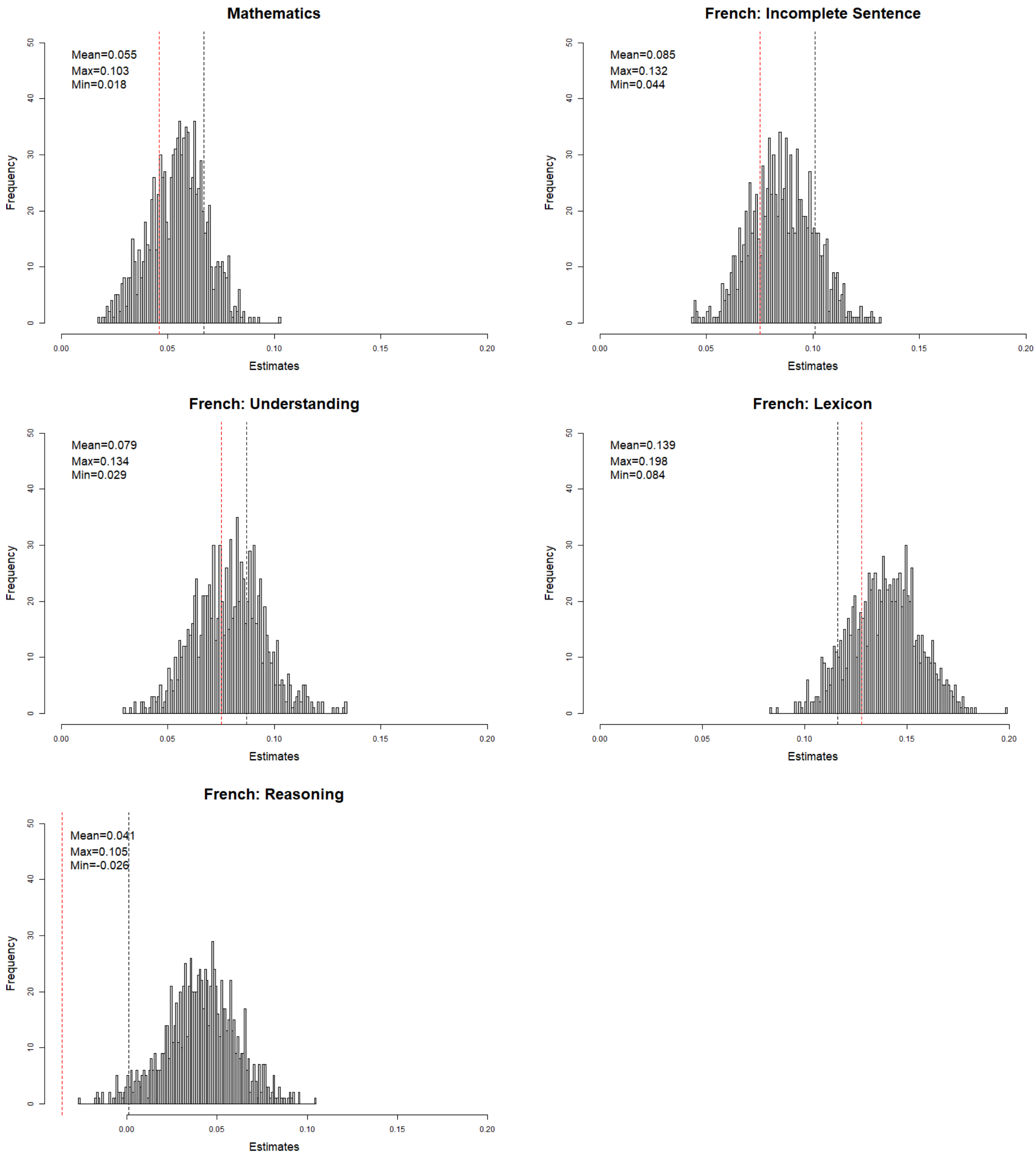

Figure 5: Histograms of 1000 realisations of the genetic matching estimates for the Difference in grade in Brevet des College exam in 2011 with a population size of 10,000 when matching on all covariates (fifth column of Table 11). The black line represents the estimate from genetic matching shown in Table 11, and the red line is the propensity score estimate shown in Table 17.

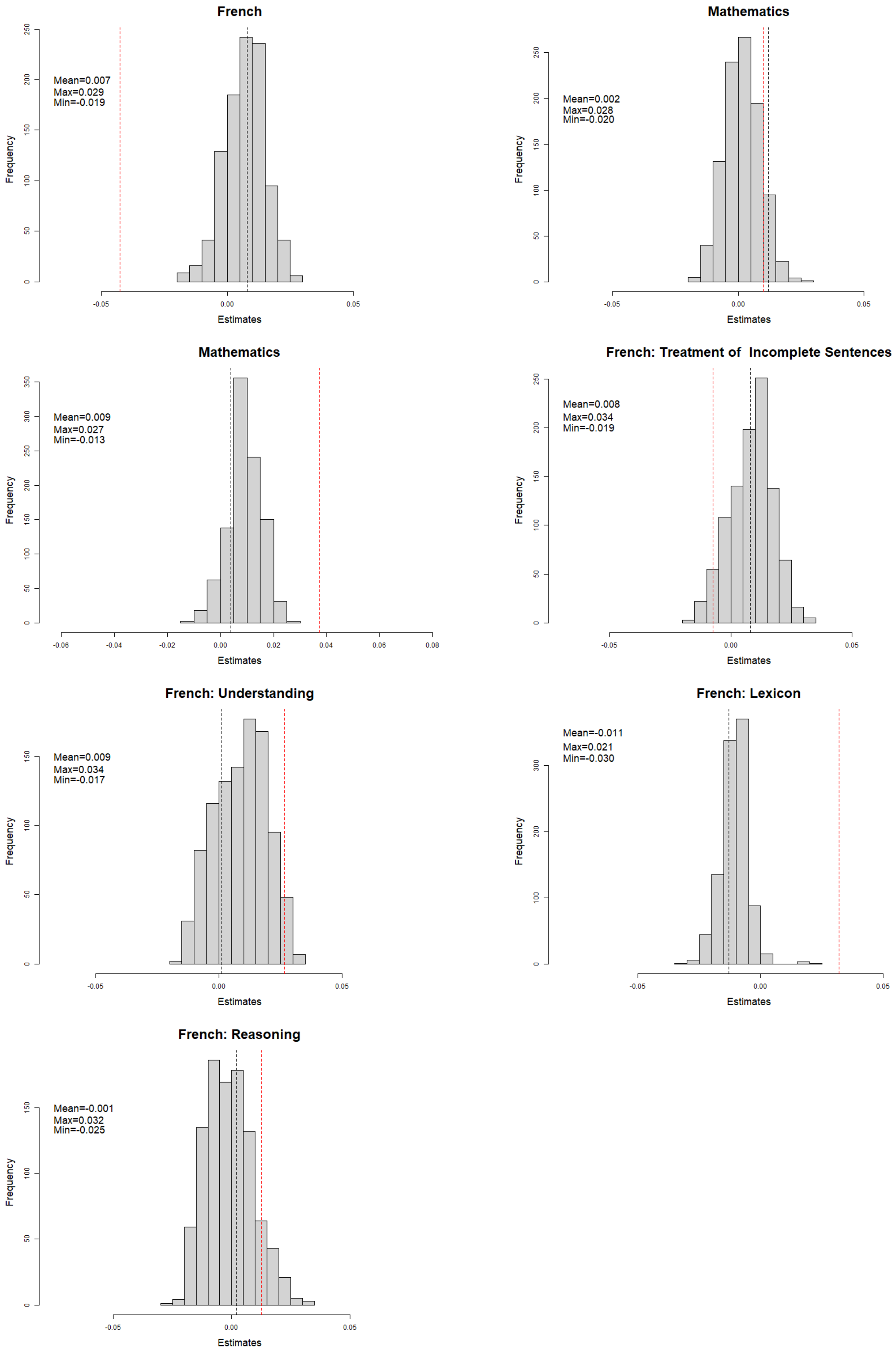

Figure 6: Histograms of balance checks for 1000 realisations of the genetic matching with a population size of 10,000 when matching on all covariates (fifth column of Table 12). The black line represents the estimate from genetic matching shown in Table 12, and the red line is the propensity score estimate shown in Table 18.

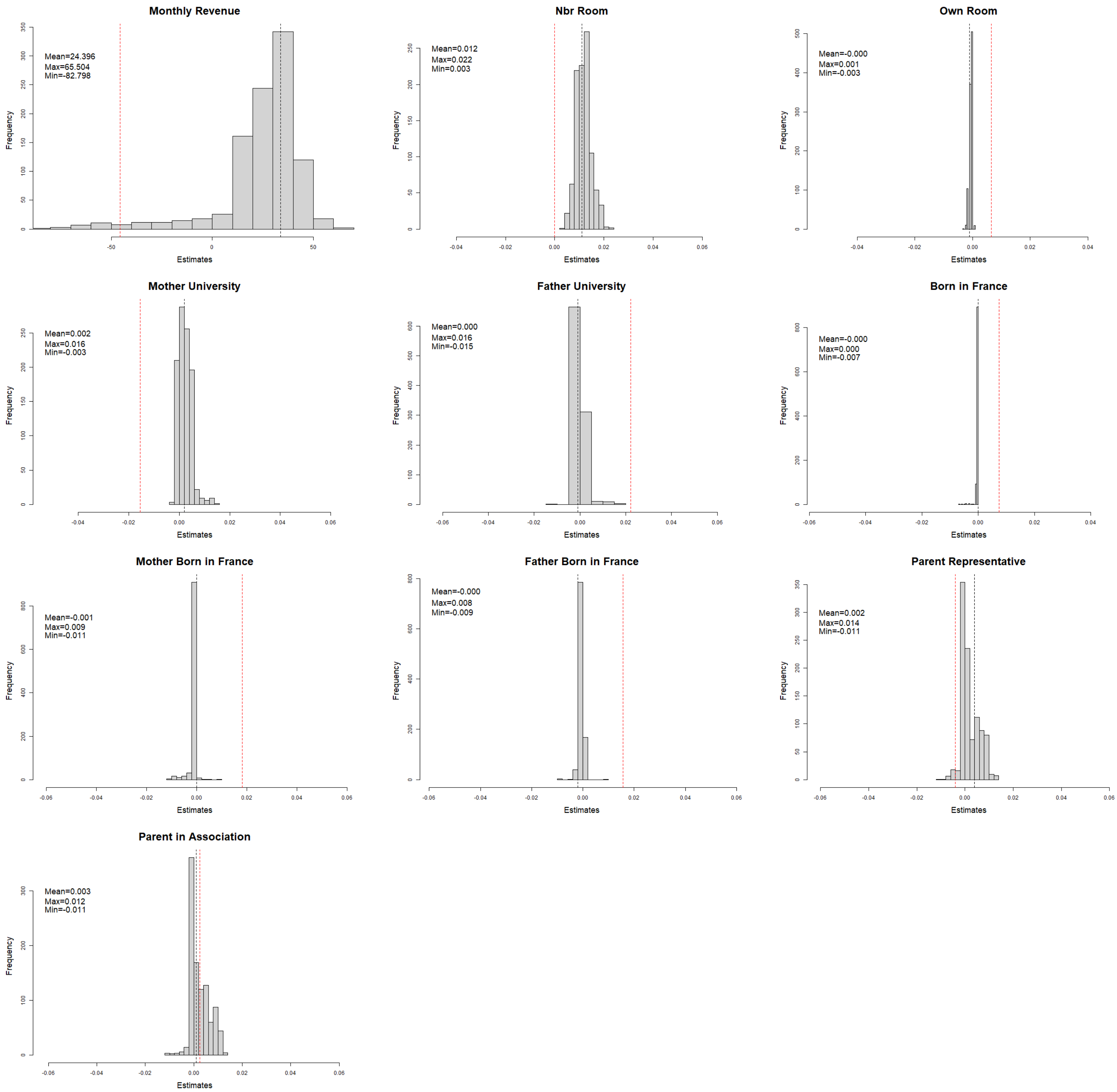

Figure 7: Histograms of balance checks for 1000 realisations of the genetic matching with a population size of 10,000 when matching on all covariates (fifth column of Table 13). The black line represents the estimate from genetic matching shown in Table 13, and the red line is the propensity score estimate shown in Table 18.

D Exam Scores Distributions

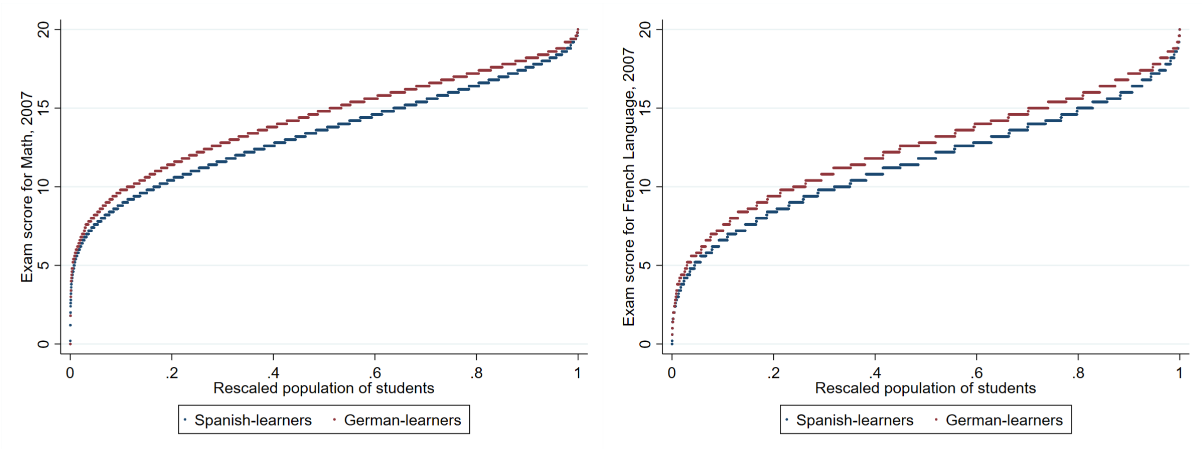

Figure 8: The graphs show the distribution of exam scores for the 2007 National Test. Populations of students are rescaled to be between zero and one; all scores are shown in ascending order; each dot represents one student, with Spanish learners in blue and German learners in red. The left panel depicts the distribution of exam scores in Mathematics. The right panel depicts the distribution of exam scores in a French language course.

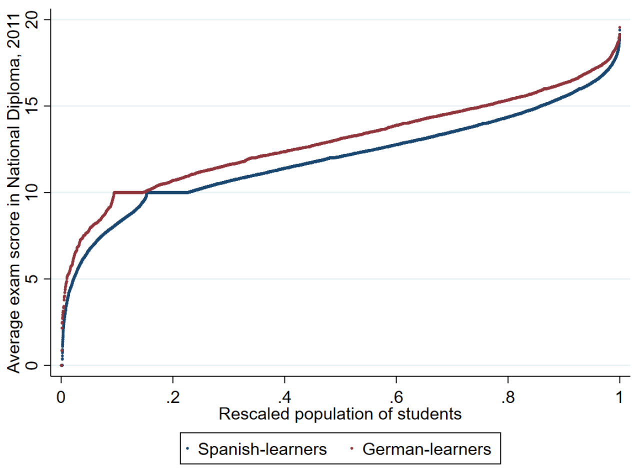

Figure 9: The graphs show the distribution of exam scores for the National Diploma. Populations of students are rescaled to be between zero and one, all scores are shown in ascending order, and each dot represents one student, with Spanish learners in blue colour and German learners in red colour.

E Genetic matching estimates

Table 1, 2, and 3 show the results for the Brevet des Colleges grades and Table 4 shows the results for the standardised tests. We performed five different matchings for each outcome variable, where we gradually included more variables in the matching. The fifth column of the tables does the matching, taking into account the 2007 National Test, the 2008 Special Test, the socioeconomic characteristics, and the parents' involvement. The more variables are included in the matching, the more credible the matching is. Therefore, the last column in each table shows the most credible estimates.

Table 1: Difference in grade in Brevet des College exam in 2011 estimated with Genetic Matching with a population size of 10,000.

	Overall Average (Cont. and Exam)				
German	0.122***	0.087***	0.080**	0.055†	0.102***
p-value	(< 0.001)	(< 0.001)	(0.003)	(0.073)	(< 0.001)
N	17950	17039	9047	7014	6980
	Overall Average (Cont.)				
German	0.116***	0.091***	0.111***	0.106***	0.108***
p-value	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
N	17950	17039	9047	7014	6980
	Mathematics (Cont.)				
German	0.078***	0.047*	0.056†	0.010	0.022
p-value	(< 0.001)	(0.023)	(0.053)	(0.740)	(0.484)
N	17914	17007	9032	7007	6973
	Mathematics (Exam)				
German	0.087***	0.045*	0.083**	0.048	0.062*
p-value	(< 0.001)	(0.022)	(0.002)	(0.109)	(0.039)
N	17645	16771	8921	6946	6913
	French (Cont.)				
German	0.113***	0.083***	0.072**	0.071*	0.075*
p-value	(< 0.001)	(< 0.001)	(0.008)	(0.020)	(0.013)
N	17915	17008	9032	7005	6971
	French (Exam)				
German	0.061***	0.032	0.066*	0.057†	0.091**
p-value	(< 0.001)	(0.104)	(0.011)	(0.058)	(0.002)
N	17675	16796	8929	6951	6918
	French (Dict.)				
German	0.050**	0.030	0.054*	0.059*	0.085**
p-value	(0.001)	(0.110)	(0.031)	(0.048)	(0.003)
N	17669	16793	8928	6950	6917
	French (Essay)				
German	0.064**	0.048†	0.058†	0.075†	0.080*
p-value	(0.002)	(0.062)	(0.097)	(0.072)	(0.043)
N	17661	16786	8927	6950	6917
National test	✓	✓	✓	✓	✓
Special test	✗	✓	✓	✓	✓
Family income	✗	✗	✓	✓	✓
Other socioeconomic	✗	✗	✗	✓	✓
Parent's involvement	✗	✗	✗	✗	✓

Table 2: Difference in grade in Brevet des College exam in 2011 estimated with Genetic Matching with a population size of 10,000.

	History and Geography (Exam)				
German	0.332***	0.314***	0.314***	0.376***	0.386***
p-value	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
N	17261	16385	8683	6739	6706
	History and Geography (Cont.)				
German	0.110***	0.078***	0.103***	0.072*	0.119***
p-value	(< 0.001)	(< 0.001)	(< 0.001)	(0.037)	(< 0.001)
N	17633	16760	8917	6943	6910
	Physics and Chemistry (Cont.)				
German	0.105***	0.073***	0.146***	0.077*	0.054
p-value	(< 0.001)	(< 0.001)	(< 0.001)	(0.019)	(0.115)
N	17902	16996	9029	7004	6970
	Biology (Cont.)				
German	0.127***	0.120***	0.092**	0.074*	0.080*
p-value	(< 0.001)	(< 0.001)	(0.002)	(0.030)	(0.014)
N	17902	16996	9027	7001	6967
	Technology (Cont.)				
German	0.106***	0.071**	0.074*	0.055	0.027
p-value	(< 0.001)	(0.003)	(0.021)	(0.116)	(0.434)
N	17872	16965	9012	6992	6959
National test	✓	✓	✓	✓	✓
Special test	✗	✓	✓	✓	✓
Family income	✗	✗	✓	✓	✓
Other socioeconomic	✗	✗	✗	✓	✓
Parent's involvement	✗	✗	✗	✗	✓

Table 3: Difference in grade in Brevet des College exam in 2011 estimated with Genetic Matching with a population size of 10,000.

			Plastic Art (Cont.)		
German	0.080***	0.058*	0.054	0.033	0.063†
p-value	(< 0.001)	(0.023)	(0.118)	(0.380)	(0.089)
N	17866	16961	9012	6991	6957
			Music (Cont.)		
German	0.130***	0.108***	0.134***	0.146***	0.103**
p-value	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.004)
N	17800	16901	8982	6968	6934
			Sports (Cont.)		
German	−0.006	−0.005	−0.028	−0.041	−0.035
p-value	(0.783)	(0.843)	(0.429)	(0.283)	(0.370)
N	17644	16759	8911	6918	6884
			Civic Education (Cont.)		
German	0.085***	0.056*	0.077*	0.119***	0.102**
p-value	(< 0.001)	(0.020)	(0.019)	(< 0.001)	(0.005)
N	15401	14616	7763	6034	6006
			Conduct Mark (Cont.)		
German	0.102***	0.077***	0.079*	0.070*	0.093**
p-value	(< 0.001)	(< 0.001)	(0.016)	(0.048)	(0.009)
N	17917	17010	9035	7006	6972
National test	✓	✓	✓	✓	✓
Special test	✗	✓	✓	✓	✓
Family income	✗	✗	✓	✓	✓
Other socioeconomic	✗	✗	✗	✓	✓
Parent's involvement	✗	✗	✗	✗	✓

Table 4: Difference in grade in Special Test in 2011 estimated with Genetic Matching with a population size of 10,000.

			Mathematics		
German	0.072***	0.016	0.053**	0.032	0.067**
p-value	(< 0.001)	(0.261)	(0.007)	(0.194)	(0.004)
N	15617	14916	7970	6258	6229
			French: Incomplete Sentence		
German	0.102***	0.054**	0.072**	0.090**	0.101***
p-value	(< 0.001)	(0.004)	(0.004)	(0.002)	(< 0.001)
N	15597	14900	7958	6251	6222
			French: Understanding		
German	0.069***	0.052*	0.060*	0.070*	0.087**
p-value	(< 0.001)	(0.014)	(0.043)	(0.032)	(0.008)
N	15587	14887	7951	6242	6214
			French: Lexicon		
German	0.125***	0.077***	0.109***	0.130***	0.116***
p-value	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
N	15623	14920	7968	6255	6227
			French: Reasoning		
German	0.033†	0.001	0.031	−0.009	0.001
p-value	(0.081)	(0.972)	(0.296)	(0.800)	(0.974)
N	15564	14867	7932	6231	6203
National test	✓	✓	✓	✓	✓
Special test	✗	✓	✓	✓	✓
Family income	✗	✗	✓	✓	✓
Other socioeconomic	✗	✗	✗	✓	✓
Parent's involvement	✗	✗	✗	✗	✓

F Genetic matching balancing check

Table 5: Balance check with genetic Matching with a population size of 10,000.

	National Evaluation 2007				
	French				
Difference	-0.000	0.001	0.004	0.006	0.008
t-test p-value	(0.959)	(0.822)	(0.672)	(0.698)	(0.524)
KS stat	0.003	0.019	0.019	0.025	0.024
KS bootstrapped p-value	(0.998)	(0.491)	(0.813)	(0.674)	(0.748)
	Mathematics				
Difference	0.000	-0.002	0.004	0.007	0.012
t-test p-value	(1.000)	(0.803)	(0.621)	(0.510)	(0.255)
KS stat	0.000	0.017	0.026	0.028	0.030
KS bootstrapped p-value	(1.000)	(0.670)	(0.539)	(0.624)	(0.564)
	Special Test 2008				
	Mathematics				
Difference	0.073	-0.001	0.003	0.011	0.004
t-test p-value	($< 0.001^{***}$)	(0.822)	(0.666)	(0.524)	(0.679)
KS stat	0.043	0.014	0.018	0.030	0.020
KS bootstrapped p-value	($< 0.001^{***}$)	(0.785)	(0.860)	(0.483)	(0.869)
	French: Treatment of Incomplete Sentences				
Difference	0.064	0.000	0.002	0.000	0.008
t-test p-value	(0.001^{**})	(0.904)	(0.697)	(0.988)	(0.652)
KS stat	0.044	0.009	0.014	0.029	0.028
KS bootstrapped p-value	($< 0.001^{***}$)	(0.979)	(0.939)	(0.459)	(0.483)
	French: Understanding				
Difference	0.068	0.002	-0.002	0.006	0.001
t-test p-value	(0.001^{**})	(0.913)	(0.920)	(0.764)	(0.955)
KS stat	0.038	0.013	0.016	0.029	0.021
KS bootstrapped p-value	($< 0.001^{***}$)	(0.724)	(0.854)	(0.391)	(0.641)
	French: Lexicon				
Difference	0.042	-0.001	-0.011	-0.008	-0.013
t-test p-value	($0.067^{\dagger}$)	(0.679)	(0.612)	(0.653)	(0.353)
KS stat	0.025	0.013	0.027	0.032	0.038
KS bootstrapped p-value	($< 0.001^{***}$)	(0.655)	(0.264)	(0.262)	(0.121)
	French: Reasoning				
Difference	0.065	-0.001	0.003	0.004	0.002
t-test p-value	(0.003^{**})	(0.795)	(0.638)	(0.709)	(0.852)
KS stat	0.051	0.007	0.027	0.024	0.029
KS bootstrapped p-value	($< 0.001^{***}$)	(0.999)	(0.401)	(0.645)	(0.463)
National Test	✓	✓	✓	✓	✓
Special Test	✗	✓	✓	✓	✓
Parent's Monthly Revenue	✗	✗	✓	✓	✓
Other socioeconomic	✗	✗	✗	✓	✓
Parent's involvement	✗	✗	✗	✗	✓

Table 6: Balance check with genetic Matching with a population size of 10,000.

			Socioeconomic		
			Monthly' Revenue		
Difference	34.788	-78.023	4.223	9.087	33.812
t-test p-value	(0.620)	(0.252)	(0.682)	(0.638)	(0.331)
KS stat	0.020	0.039	0.018	0.032	0.039
KS bootstrapped p-value	(0.065†)	(0.171)	(0.905)	(0.479)	(0.242)
			Nbr Room		
Difference	0.162	0.169	0.157	0.004	0.011
t-test p-value	($< 0.001^{***}$)	($< 0.001^{***}$)	(0.001^{**})	(0.692)	(0.389)
KS stat	0.042	0.035	0.041	0.010	0.017
KS bootstrapped p-value	($< 0.001^{***}$)	(0.012^{*})	(0.035^{*})	(0.969)	(0.680)
			Own Room		
Difference	-0.010	-0.007	-0.001	0.000	-0.001
t-test p-value	(0.328)	(0.516)	(0.926)	(1.000)	(0.564)
			Mother University		
Difference	0.013	0.008	-0.001	0.001	0.002
t-test p-value	(0.279)	(0.528)	(0.947)	(0.655)	(0.480)
			Father University		
Difference	0.023	0.026	0.016	0.000	-0.001
t-test p-value	(0.067†)	(0.040^{*})	(0.290)	(1.000)	(0.564)
			Born in France		
Difference	-0.002	-0.007	-0.006	0.000	0.000
t-test p-value	(0.566)	(0.103)	(0.323)	(1.000)	(1.000)
			Mother Born in France		
Difference	-0.004	-0.011	-0.008	0.000	0.000
t-test p-value	(0.671)	(0.214)	(0.516)	(1.000)	(1.000)
			Father Born in France		
Difference	0.002	0.002	-0.007	0.000	-0.002
t-test p-value	(0.861)	(0.869)	(0.615)	(1.000)	(0.480)
			Parent Representative		
Difference	0.024	0.017	-0.012	0.002	0.004
t-test p-value	(0.005^{**})	(0.052†)	(0.327)	(0.907)	(0.662)
			Parent in Association		
Difference	0.028	0.025	0.007	-0.011	0.001
t-test p-value	(0.004^{**})	(0.015^{*})	(0.596)	(0.475)	(0.706)
National Test	✓	✓	✓	✓	✓
Special Test	✗	✓	✓	✓	✓
Parent's Monthly Revenue	✗	✗	✓	✓	✓
Other socioeconomic	✗	✗	✗	✓	✓
Parent's involvement	✗	✗	✗	✗	✓

G Survey Questions

Below are the survey questions in French and English. For each question, we randomise the order in which choices are displayed. We applied the same randomisation order for Q1 and Q3, with "Other" fixed as the last option. The same approach is applied for Q2 and Q4. Additionally, the options for Q5 and Q6 are randomised (in the same order across both questions), and we also randomised the order of the questions so that some respondents will answer Q6 before Q5. Lastly, we applied a reverse order for Q7 options to avoid having "Yes, No, and Yes" as the option order.

YouGov performs quality control on the responses it receives. Responses are removed if suspicious activity is detected. For example, duplicate responses are identified and removed based on a combination of demographic information, answers to quality control questions, and IP addresses. Additionally, respondents who complete the survey unusually quickly are excluded. More information about YouGov's methodology can be found here: https://today.yougov.com/about/panel-methodology.

G.1 Questions in French

1. Au collège, quelle première langue vivante (LV1) avez-vous étudiée ?
   (a) Anglais
   (b) Allemand
   (c) Espagnol
   (d) Italien
   (e) Autre (précisez):
2. Quelles ont été les motivations pour choisir cette première langue? (allow multiple choice)
   (a) Pour intégrer une classe de bon niveau et être avec des camarades motivés ou performants
   (b) Parce que cette langue me semblait plus intéressante et/ou utile
   (c) Parce que cette langue me semblait plus facile à apprendre
   (d) Pour montrer mes compétences et mon intelligence
   (e) Autre (précisez):
3. Au collège, quelle seconde langue vivante (LV2) avez-vous étudiée ?
   (a) Espagnol
   (b) Allemand
   (c) Anglais
   (d) Italien
   (e) Autre (précisez):
4. Quelles ont été les motivations pour choisir cette seconde langue? (allow multiple choice)

(a) Pour intégrer une classe de bon niveau et être avec des camarades motivés ou performants
(b) Parce que cette langue me semblait plus intéressante et/ou utile
(c) Parce que cette langue me semblait plus facile à apprendre
(d) Pour montrer mes compétences et mon intelligence
(e) Autre (précisez):

5. Selon vous, l'espagnol ou l'allemand est-il plus difficile à étudié?
   (a) Espagnol
   (b) Allemand

6. Selon vous, l'espagnol ou l'allemand est-il plus utile à parler?
   (a) Espagnol
   (b) Allemand

7. Vos parents ont-ils joué un rôle actif dans le choix de votre seconde langue vivante?
   (a) Oui, ils m'ont fortement encouragé(e) ou ont insisté pour une langue en particulier
   (b) Oui, ils m'ont donné des conseils, mais le choix final m'appartenait
   (c) Non, ils n'ont pas influencé ma décision

G.2 Questions in English

1. Which first foreign language (LV1) did you study in middle school?
   (a) English
   (b) German
   (c) Spanish
   (d) Italian
   (e) Other (specify):

2. What were your motivations for choosing this first language? (allow multiple choice)
   (a) To join a high-achieving class and be with motivated or high-performing classmates
   (b) Because the language seemed more interesting and/or useful
   (c) Because the language seemed easier to learn
   (d) To show my skills and intelligence
   (e) Other (please specify):

3. Which second foreign language (LV2) did you study in middle school?
   (a) Spanish

(b) German
(c) English
(d) Italian
(e) Other (please specify):

4. What were your motivations for choosing this second language? (allow multiple choice)

   (a) To join a high-achieving class and be with motivated or high-performing classmates
   (b) Because the language seemed more interesting and/or useful
   (c) Because the language seemed easier to learn
   (d) To show my skills and intelligence
   (e) Other (please specify):

5. In your opinion, which language is harder to learn?

   (a) Spanish
   (b) German

6. In your opinion, which language is more useful to speak?

   (a) Spanish
   (b) German

7. Did your parents play an active role in the choice of your second foreign language?

   (a) Yes, they strongly encouraged or insisted on a particular language
   (b) Yes, they gave advice, but the final choice was mine
   (c) No, they did not influence my decision

H Comparative Statics for an Intermediate Equilibrium Threshold

Consider an intermediate equilibrium threshold, $a' \in A^* \backslash \{\underline{a}^*, \bar{a}^*\}$. We want to prove the claim below:

Claim. *Suppose that* $\frac{\partial \Delta E[u_i(a;\lambda)]}{\partial a} \neq 0$ *at* $a = a'$. *If* $\frac{\partial \Delta E[\mathcal{P}_i(a=a')]}{\partial a} \geq 0$, *then* $\frac{da'(\lambda)}{d\lambda} \geq 0$.

Proof. Suppose that $\frac{\partial \Delta E[\mathcal{P}_i(a=a')]}{\partial a} \geq 0$. Now suppose that $\frac{da'(\lambda)}{d\lambda} < 0$. By the implicit function theorem, we have

$$\frac{da'(\lambda)}{d\lambda} = -\frac{\partial \Delta E[u_i(a';\lambda)]/\partial \lambda}{\partial \Delta E[u_i(a';\lambda)]/\partial a} < 0.$$

Note that, since $\frac{\partial \Delta E[u_i(a';\lambda)]}{\partial \lambda} = -\frac{\partial \Delta c(a',\lambda)}{\partial \lambda} < 0$, we have

$$\frac{\partial \Delta E[u_i(a';\lambda)]}{\partial a} = \frac{\partial \Delta E[\mathcal{P}_i(a')]}{\partial a} - \frac{\partial \Delta c(a',\lambda)}{\partial a} < 0.$$

Thus, we have

$$0 \leq \frac{\partial \Delta E[\mathcal{P}_i(a')]}{\partial a} < \frac{\partial \Delta c(a',\lambda)}{\partial a} < 0,$$

which is a contradiction. Hence, $\frac{da'(\lambda)}{d\lambda} \geq 0$. □