ON PRIVATE UNOBSERVED RETURNS TO INTERNATIONAL MIGRATION IN A COUPLE

ROMUALD MÉANGO AND TILL NIKOLKA

Abstract. This paper examines the private unobserved migration propensity of married individuals using bounds to circumvent the issue of partial observability. Applied to the population of Danish couples aged between 25 to 39, this approach leads to two main results. First, we find convincing evidence that married individuals differ from single individuals in their migration propensity even after controlling for their observable characteristics. Second, after assessing the relative importance of male and female partners' characteristics in the decision to emigrate, we cannot reject the hypothesis that both partners' observed characteristics are equally weighted in the migration decision.

Keywords: bounds; couple; migration; partial observability; selection.

JEL Codes: F22; J12; J16; J24

1. Introduction

The conflict in a couple raised by divergent private returns to migration is of high relevance for societies that have experienced a dramatic change in the female labor force participation. This conflict is even more relevant when the interest is in international migration decisions, which are associated with large potential returns and equally significant losses. Mincer (1978) first coined the terms *tied stayers* and *tied movers* to describe couple members who, in order to remain in the couple, choose a different location than the one that maximizes their private returns. Although conceptually easy to understand, tied mobility is difficult to identify in many datasets, for once an individual is married, observation of the couple migration behavior conveys little information on the private incentives of each spouse. The issue is one of "partial

Date: July 24, 2014. Ifo Institute, Munich. Financial support by the Leibniz Association (SAW-2012-ifo-3) is gratefully acknowledged. The authors are thankful to Stéphane Bonhomme, Johannes Köenen, Panu Poutvaara, and Ismael Mourifié for helpful discussions. All remaining errors are ours.
Correspondence address: meango@ifo.de .

observability" (Poirier, 1980). From observing a migration event, we can infer that one of the partners has significant migration incentives; however, we are not able to identify the partners' own private incentives, that is, what their decision would be in a counterfactual scenario uninfluenced by family ties[1].

Identifying private incentives is particularly relevant when trying to understand the higher responsiveness of couples to male's career opportunities compared to those of female's, a result supported by a large body of research[2]. An important question is whether this asymmetry could be due to gender differences in the private returns to migration, that are unrelated to marriage (he has an unobserved taste for migration, she does not). To perform a proper comparison of migration returns by gender, one should observe the partners' private incentives, something not typically possible. To circumvent the problem of partial observability, Tenn (2010) uses the singles in the population as a counterfactual group to perform a variance decomposition of the couple migration decision, assuming that singles and married individuals do not differ on unobservable characteristics influencing both marriage and migration decisions. However, the underlying assumption of no selectivity in marriage is arguable as recognized by, for example, Gemici (2007) and Sorenson and Dahl (2012). Although this intuition could be traced back to Mincer (1978), to the best of our knowledge, there exists no empirical study that convincingly assess the claim that differences in unobservables between married and singles matter when it comes to migration decisions, to the point that singles are poor counterfactuals for married.

We pursue two objectives in this paper. The first is to test whether individuals in couples differ from singles on unobservable characteristics that also influence their migration decision. Answering this question is crucial to understanding whether singles' migration propensity provides a reliable counterfactual for their married counterparts' unobserved migration propensity. Our results suggest that it does not, at least not in the female population. The second objective is to properly assess the relative

[1]The issue of partial observability is even more acute than in the standard Poirier model, since potential compensation for one party's loss can be achieved by the other party; disagreement over the migration decision is therefore neither a proper subset of the migration event, nor a proper subset of the staying event.

[2]See for example Duncan and Perrucci (1976); Nivalainen (2004); Taylor (2007) and Rabe (2011).

importance of male's and female's characteristics in the couple's migration decision, controlling for gender differences in the return to migration. Given the inappropriateness of singles to serve as proxies, we propose using observed decisions of couples to construct bounds on the private migration incentives. We then assess the relative importance of each members observed characteristics in the couple migration decision. To this end we derive bounds on the ratio of the variance of the mean return to migration.

As an important contribution, we base our assumptions on economic theory and stylized facts taken from the literature about couple migration. Specifically, we motivate the use of the couple's migration propensity as a lower bound for the migration propensity of its members. In addition, we motivate the use of the migration propensity of separated couples as an upper bound of the unobserved migration probability. Compared to the bounds one obtains following the proposals of Manski (1995) and Manski and Pepper (2000) (henceforth MMP), the proposed bounds are considerably tighter.

For our empirical analysis we use data from the full Danish population register for the years 1990 to 2005. We observe couples and singles based on their cohabiting or marriage status in Denmark. As registration is compulsory in the case of emigration as well as for return migration, we have a reliable source for the number of long-term emigration events. We perform separate analysis for males and females. Aside form the richness of the dataset, the case of Denmark is in itself interesting in many respects. First, as it is true of for other Scandinavian countries, Denmark is perceived as one the most gender-equal societies in the world. Females are in average better educated than males and their labor force participation has been over 70% for several decades. Nonetheless, the gender income gap has not been completely eliminated. Thus, Denmark's situation may provide useful insight for other countries that are experiencing closing income gaps between men and women and where female labor force participation is increasing. Second, while to some degree, Danish single females appear to be more mobile than Danish single males, Danish couples' international migration seems to respond more to male characteristics than to female ones. This reversal of the gender asymmetry suggests either a selection into couple or a "traditional" decision-making process in couples. The distinction between the two is an empirical question.

We first test whether married individuals differ from single individuals in their migration propensity even after controlling for their observable characteristics. The test consists in assessing whether the migration probability of a comparable single lies between the bounds we devise. From this procedure, we conclude that married individuals are selected on unobservable characteristics, especially females. We argue for two sources of selection. First, some unobserved traits are important to the international mobility of an individual and to the formation of a couple. Second, a single with a high return to international migration might delay becoming part of a couple in order to pursue a migration opportunity.

Then, we assess the relative importance of male and female partners' characteristics in the migration decision of Danish couples and find that we cannot reject the hypothesis that both partners' observed characteristics are equally weighted in the migration decision. This result contradicts the results achieved using Tenn's procedure.

Marriage is associated with a large number of outcomes, including earnings, physical health, and children. Selectivity into marriage is a major concern in this body of research. For a comprehensive survey of this body of research, see Ribar (2004). This literature relies heavily either on exogenous variation in attractiveness using randomized experiments, changes in laws or the so-called shot-gun marriage, and/or, alternatively, on strong structural assumptions. A major comparative advantage of the approach proposed here is the relative weakness of the assumptions and their transparency. Recent work from Tano, Westerlund, Nakosteen, and Zimmer (2014) also links selectivity issues in the formation of couples to the location choice of agents. Selectivity occurs in the formation of power couples (couples where both partners are highly educated) on the same unobservable variables that influence the decision to locate in large metropolitan areas. Again, the assumptions used here are weaker than theirs.

Manski, Sandefur, McLanahan, and Powers (1992) is one of the early studies to use non-parametric bounds to interpret the association between family structure and high school graduation found among respondents in the National Longitudinal Study of Youth. Angrist, Bettinger, and Kremer (2004) use bounds to identify the effect of school vouchers on test scores, where the selection problem arises because not all treated pupils opt to take the test. This paper is closely related to that of Blundell, Gosling, Ichimura, and Meghir (2007) who use bounds to test for selection into unemployment. Their work also discusses the use of bounds in a wide range of economic applications (see Tamer, 2010, for a recent survey). The construction of the bounds used here takes advantage of stylized facts in the migration literature, allowing for a significant reduction of the bounds proposed by MMP. To the best of our knowledge, using the migration behavior of couples and separated couples to bound the private migration propensity of couples is a first in the literature.

The migration literature consistently documents higher responsiveness of couples to male's career opportunities than to female's and this is true in the Nordic countries, too: on internal migration in Finland, Nivalainen (2004), on internal migration in Sweden, Axelsson and Westerlund (1998), on internal migration in Denmark Sorenson and Dahl (2012) and on international migration from Denmark, Munk, Junge, and Poutvaara (2013). Further evidence on the primacy of the male's career can be found in numerous studies finding that labor market outcomes are in general better for men than for women following household migration (Boyle, Cooke, Halfacree, and Smith, 1999; Jacobsen and Levin, 1997; Åström and Westerlund, 2009). As mentioned above, this paper studies the relative contribution of male's characteristics to female characteristics in the migration decision, as proposed by Tenn (2010). However, Tenn uses singles are used as counterfactuals to control for possible gender differences in the migration propensity whereas our methodological contribution is to instead use bounds on this quantity.

The bounds we propose belong to the class of intersection bounds. To conduct inference, we follow the methodology advised by Chernozhukov, Lee, and Rosen (2013). Their work, however, does not contain a proposal for constructing a valid confidence for the unconditional monotone instrumental variable (MIV) case, as devised by Manski and Pepper (2000). As another contribution of possibly independent interest, we propose a two stage procedure to retrieve valid confidence regions for the unconditional MIV case, when the monotone instrument has a finite support. We briefly explore the finite sample properties of this procedure in Appendix D.

The rest of the paper is organized as follows. Section 2 introduces our bounds on the *private* migration propensity of an individual in the context of the selection test we wish to perform. We briefly discuss the sharp bounds of MMP before introducing the stylized facts of the migration literature, the use of which improves the bounds. Section 3 describes the dataset, some summary statistics, and the inference methodology, as well as the implicit bounds on the variance decomposition ratio. This section also contains the proposal for constructing bounds from an unconditional MIV. Section 4 performs successively, the tests of selection under alternative assumptions, followed by the variance decomposition exercise. Finally, Section 5 concludes the discussion.

2. A Test of Selection

Before discussing the test procedure,we need to make clear our definition of individual *private incentives* to migrate. By private incentives, we mean the gains (or losses) an individual would experience from migration if (s)he were to make this decision for himself or herself alone. In other words, our interest is in the decision an agent would make had (s)he no family tie. It is crucial not to confuse these private incentives with the incentives an individual would have had (s)he always remained single. Our view is that marriage shapes individual characteristics. The relevant point of comparison when discussing issues such as family decision-making processes or tied mobility is the private incentives of an individual at the point of time the researcher observes him/her.

This definition makes it clear why singles might not be good proxies for their married counterparts. There are at least three ways the two groups might differ: (i) selection on unobserved heterogeneity, (ii) sorting (out of marriage) of *migration-prone* individuals, and (iii) a difference in investment in unobserved skills.

First, there might be latent abilities and unobserved traits that are important both to the mobility of an individual and to the formation of a couple. For example, according to Alfred Marshall's *Principles of Economics*, migration to large labor market areas has been associated with the "most enterprising, the highly gifted, those with highest physique and strongest character"[3].

Second, in an environment where household surplus is limited, returns to migration are in general negative, and family ties decrease migration propensity (see below), agents with low or no incentive to migrate will be more likely to mate than agents with high returns to migration. The rationale for this idea is that the latter group is less likely to find a match who will either be willing to migrate to the same destination or able to generate sufficient household surplus to compensate the losses from not migrating[4]. Based on this reasoning, we might therefore hypothesize a sorting of *migration-prone* individuals out of marriage.

The third way singles and couple members might differ in their migration incentive is that both groups might invest differently in specific skills that are more valued in the event of migration. A comparative advantage of marriage is that it allows for task specialization in home production and market work (Waite and Gallagher, 2002). As people specialize, they may increase their skills through experience or training and become even more productive in these activities over time[5] (Ribar, 2004).

[3]Quoted in Greenwood (1997). With respect to international migration from Denmark, Kauppinen, Borjas, and Poutvaara (2013) document a first-order stochastic dominance of the distribution of migrants' unobservable characteristics on the distribution of non-migrants' unobservable characteristics. There are good reasons to think that these characteristics make someone more attractive on the marriage market. Gautier, Svarer, and Teulings (2010), studying a Danish cohort born in the 1950's and 1960's, find that their measure of economic attractiveness on the marriage market predicts to some extent the propensity of an individual to be highly-educated and, hence, his/her choice of location. For a Swedish cohort born in the 1980's, Tano, Westerlund, Nakosteen, and Zimmer (2014) show that both the formation of a power-couple and the choice of location in a large metropolitan area are positively influenced by individual unobserved heterogeneity.

[4]Greenwood, Guner, Kocharkov, and Santos (2012) find empirical evidence that technological change, the change in educational premiums, and the closing gender wage gap together allow highly educated singles to be more selective in choosing a partner,leading to a reduction in household size and a surge in assortative mating in the United States over the past decades.

[5]A high degree of specialization enables the agents to achieve a higher level of total surplus (Gemici and Laufer, 2011). Guler, Guvenen, and Violante (2012) propose a joint-search model where couples' members take advantage of living together to achieve better job match and higher wages through a "single bread-winner" cycle.

Note that these mechanisms might work in opposite directions. The second mechanism implies that some subgroup of singles will have higher migration returns than the average married person, even after controlling for observable characteristics. The first and third potential differences give rise to the expectation that a subgroup of married individuals has higher private migration incentives as their single counterparts, even after controlling on observable characteristics. We now turn to what we refer as the *test of selection*.

2.1. **Notation.** Let Y_i be the migration decision of an individual i, where Y_i is a binary equal to 1 when choosing to migrate would maximize her private utility. Denote M_i the marital status of an individual i. Assume that M_i can take one of three values, c for belonging to a couple, s for being single, and d for an individual who divorced[6]. Y_i is fully observed when $M_i = s$. However, the migration event we observe when $M_i = c, d$, say, Y_{ci} and Y_{di}, might not entirely reflect the private migration incentives. X_i denotes the random variable summarizing the observable characteristics of i, here, educational attainment and age.

Our interest is in testing the assumption:

$$H_0 : \mathbb{P}(y|M = c, X) = \mathbb{P}(y|M = s, X) = \mathbb{P}(y|X, M \in \{s, c\}) \tag{2.1}$$

For ease of exposition, we will denote $F(y|X) = \mathbb{P}(y|X, M \in \{s, c\})$.

2.2. **Bounds from Manski (1995) and Manski and Pepper (2000).** As noted earlier the conditional distribution of migration as a single in the population is only partially observed. The decision an agent would have taken to maximize her private utility can be observed as long the person remains single. However, once married, the migration observed decision pertains to the couple as a unit, and individual private views toward migration are not observable. To be more precise, note that in the following

$$F(y|X) = F(y|M = s, X)\mathbb{P}(M = s|X) + F(y|M = c, X)(1 - \mathbb{P}(M = s|X))$$

$F(y|M = c, X)$ is not observed from the data. Noting that this term is a probability distribution, hence bounded between 0 and 1, Manski (1995) proposes the worst-case

[6]See Section 3.1 for the definition of these category in practice.

bounds for the migration decision:

$$F(y|M=s,X)p(X) \leq F(y|X) \leq F(y|M=s,X)p(X) + (1-p(X))$$

where $p(X)$ is a notation for $\mathbb{P}(M=s|X)$. Although intuitive, the usefulness of these bounds is limited. In particular, they do not allow assessing the existence of selection.

2.2.1. *Exclusion Restriction.* A further contribution from Manski (1995) is to show that the existence of an exclusion restriction might refine the above bounds. If Y is independent of Z conditional on X, i.e.

$$F(y|x,z) = F(y|x) \;\; \forall y, x, z$$

then the bounds to $F(Y=1|X)$ are given by:

$$\max_z F(y|M=s,X,z)p(X,z) \leq F(y|X) \leq \min_z F(y|M=s,X,z)p(X,z) + (1-p(X,z)). \tag{2.2}$$

To tighten the worst-case bounds, it is important that the instrumental variable influences the migration decision. We propose to use the marital status of both biological parents at age 25 as an excluded variable. The rationale is that parents' decision to seal a marriage contract or to dissolve one might influence the adult child in his own decision[7]. However, it is unclear how parents' marital status could affect the migration decision.

2.2.2. *Monotonicity.* An exclusion restriction might be too strong an assumption. Indeed, parental marital status (and mainly divorce) has an effect on a wide range of children outcomes in their adulthood (see Amato, 2000). Manski and Pepper (2000) propose as an alternative assumption a monotonicity restriction on a variable of interest. To be more precise, assume that the conditional migration propensity of

[7]Some of the literature suggest that marital behavior is transmitted to some extent between parents and children. Cherlin, Kiernan, and Chase-Lansdale (1995), using a British longitudinal national survey of children, show that by age 23, those whose parents divorced were more likely to cohabit and to have a child outside marriage than were those whose parents did not divorce. Amato (1996), using U.S. national longitudinal data, shows that divorce was less likely in families in which neither the husbands nor the wifes parents divorced.

an agent decreases with increasing values of a variable Z, i.e.

$$F(Y=1|x,z) \leq F(Y=1|x,z') \;\; \forall z > z' \tag{2.3}$$

For a value z_0 of Z, denote:

$$F^u(Y=1|x,z_0) \equiv \max_{z \geq z_1} F(Y=1|M=s,x,z)p(x,z) \tag{2.4}$$

and accordingly

$$F^l(Y=1|x,z_0) \equiv \min_{z \leq z_1} F(Y=1|M=s,x,z)p(x,z) + (1-p(x,z)) \tag{2.5}$$

The bounds to the conditional distribution of migration are derived by integration over the distribution of Z given $X=x$, i.e.

$$E_Z F^l(Y=1|M=s,X,Z) \leq F(Y=1|X) \leq E_Z F^u(Y=1|M=s,X,Z) \tag{2.6}$$

There exists a consensus in the migration literature that number of children is negatively associated with migration propensity. We use this variable to satisfy the monotonicity restriction.

Table 1. Worst-case bounds on $P(Y=1|M=c)$ and single migration rate in the population of female aged 30-35, by education level

Educational attainment	LB	UB	Singles
Low education	0.07%	64.77%	0.19%
Lower middle educ.	0.11%	74.60%	0.42%
Upper middle educ.	0.12%	74.40%	0.45%
High education	0.35%	68.47%	1.11%

The empirical content of the exclusion restriction and of the monotonicity condition can be used to test for the existence of selection. A major caveat of this methodology is that migration is a rare event. Table 1 computes the worst-case bounds on the population of females aged 30-35, by education group. Observe that the upper bounds are many times larger than the lower bounds. The point is that bounding above $F(Y=1|M=c,X)$ by 1 when it is suggested in the population to be very close to

0, induces a loss in informativeness of the bounds[8]. In particular, it is almost certain that $F(Y=1|M=s,X) < (1-p(x,z))$ for all z, for migration is far less common than marriage, in all sub-groups. This inadequacy of the upper bound urges us to seek alternatives bounds, since from the theory we expect the selection of migration-prone individuals out of marriage. Indeed, even an arguably good instrument would have to induce very sharp changes in the marriage decision to convey any additional information from the upper bound. We therefore make an additional assumption in order to tighten the bounds. In the following, we derive bounds on the migration probability of an agent in a couple when he is in the counterfactual situation of having no family tie, based on previous results in the literature on couple migration.

2.3. **Restrictions Specific to Couple Migration.** The previous restrictions could be used in more general contexts, albeit only for the variables satisfying the restrictions. Our goal in this section is to take advantage of specific features related to the migration decision and to the decision-making process of a couple. We introduce tighter lower and upper bounds for the unobserved probability mass $F(Y=1|M=c,X)$.

We motivate the new bounds through two stylized facts observed in the couple migration literature:

(1) Family ties are an impediment to migration, and
(2) there is a positive interaction between couple separation and individual mobility.

We discuss below the literature sustaining these two stylized facts and describe the restrictions that they impose on individuals' migration propensity. To provide some intuition on how the proposed stylized facts arise, Section 2.3.3 discusses a model of couple migration decision and the type of conditions under which the stylized facts postulated above are valid.

[8] An alternative would be to assume that migration events only concern around the 5% of the population, i.e. the migration propensity of an individual in a couple would not be higher than twice the migration propensity of the most mobile single. This approach is relatively arbitrary and we do not pursue it further.

2.3.1. *Stylized Fact 1: Family Ties as an Impediment to Migration.* Mincer (1978), Sandell (1977) and Frank (1978) all argue that marriage hinders individuals' mobility. Couples face the more complex task of reconciling two career goals in one location. Costa and Kahn (2000) link the couple's co-location problem to the high concentration of highly educated couples in metropolitan areas. Guler, Guvenen, and Violante (2012) propose a joint-search model in multiple locations that shows that the disutility of living separately restricts the number of feasible job offers for couples, hence decreasing individual mobility. Gemici (2007) estimates a structural model to quantify the decrease in interstate mobility related to family ties over the past decade in the United States. According to his estimates, family ties can decrease migration by as much as one-quarter of the initial migration probability of an individual. Finally, Munk, Junge, and Poutvaara (2013), covering the period in our analysis, find that singles in Denmark are many times more internationally mobile that their married counterpart and attributes this effect to the deterrent effect of marriage.

We therefore conjecture that on average, the migration probability of the couple to which an individual i belongs, is lower than the private migration propensity of i, that is

$$F(Y_i = 1|M_i = c, X_i) \geq \mathbb{P}(Y_{ci} = 1|M_i = c, X_i) \quad \forall X_i \tag{2.7}$$

where we denote as Y_{ci} the decision of the couple to which i belongs.

It is very important to understand that the above restriction does not rule out the possibility that some individuals are made more mobile through mating. It implies that the "average individual" is not. In assuming this, we operationalize one of the strongest stylized fact of the couple migration literature. The rationale is that tied staying is more common than tied moving, that a great job opportunity that compensate both partners losses does not happen as often as an offer that satisfies a single partner. The intuition of an economic model and stronger conditions are discussed below.

2.3.2. *Stylized fact 2: Positive Association Between Couple Separation and Migration.* Mobility and couple separation are closely linked. Note that we can distinguish separation related to a migration decision from separation following a negative shock to the household surplus. With regard to the first type, couple separation can be the result of divergent returns to migration. In particular, couples who choose separation are couples for whom migration returns are more extreme than for the general population. Consequently, conditional on the former partner not migrating, a member from a separated couple should have higher returns of migration than a similar individual in a couple. Mincer (1978) emphasizes the link between marital instability and the migration decision of a couple. Gemici (2007) attributes as much as one third of the observed divorce rate in the United States to divergent migration incentives.

The second type of couple separation, even if unrelated to previous migration opportunities, could also affect the observed migration behavior. A stream of literature related to residential mobility following separation argues that couple dissolution is a significant life event that has a positive effect on partners' mobility (Grundy, 1985; Feijten and Van Ham, 2007; Feijten and van Ham, 2013). This is because some people will move to avoid stigma, for a fresh start or to avoid contact with the ex-spouse and/or his/her family (Symon, 1990). Moves triggered by divorce, however, are restricted compared to moves triggered by other life events (Feijten and van Ham, 2013), child custody arrangements being the main restriction. Gram-Hanssen and Bech-Danielsen (2008) find that in Denmark, non-custody fathers live significantly closer to their children than childless men live to the home of their ex-partner. Since our interest is in international migration, we think that the observed migration propensity reflects split up following divergent returns to migration rather than mobility triggered by union dissolution. Note that to address the sensitivity of our results to the child custody issue, we also conduct our analysis with only childless separated couples. The results are unchanged.

Note that the term of "couple separation" is not limited to divorce or union dissolution. For example, thanks to lower costs of transportation and communication, an increasing fraction of couples decides to incur the cost of living separately, and yet maintaining their couplehood, so that both can achieve their career goals. The assumption and interpretation above still hold for couples living in different locations, since we expect that only couples where one member has very high returns to international migration will choose to forego their household surplus. Thus, for the sake of simplicity, we maintain the terminology of separated couples.

The migration probability of separated members of a couple, conditional on non-migration of the partner, should provide us with an upper bound, i.e.

$$F(Y_i = 1|M_i = c, X_i) \leq \mathbb{P}(Y_{di} = 1|M_i = d, X_i, Y_{dj} = 0) \quad \forall x_f \tag{2.8}$$

where $M_i = d$ means that i is divorced; Y_{di} (resp Y_{dj}) is the migration behavior of individual i (resp. j) that we observe following the couple separation.

Note that the restrictions expressed by Eq. (2.7) and Eq. (2.8) can be combined with the exclusion restriction and the monotonicity condition.

2.3.3. *Assessing the Stylized Facts in a Model of Couple Migration.* In order to provide some intuition on how the proposed stylized facts arise, we present a model of couple migration decision. Like Mincer (1978) we assume that both partners' gains from migration z_a and z_b follow a bivariate probability distribution. We assume that $med(z_a) \leq 0$ and $med(z_b) \leq 0$ such that the majority of the population would not emigrate even in the absence of migration costs. The partners' migration costs are denoted with $c_a \geq 0$ and $c_b \geq 0$. For simplicity, there is no uncertainty on wages abroad, i.e. the realizations of z are known to both partners when deciding on emigration. The surplus from marriage that the partners can consume when being together - at home or abroad - is denoted by h. Like in the framework proposed by Gemici (2007), assume that the migration decision follows from a bargaining process between the two partners. The partners bargain on income distribution in the family using their outside options in the home country or abroad as divorce threat points. We assume a full commitment on future transfers, or equivalently, partners experiencing losses can be compensated *ex ante.* We discuss a generalization of the framework later in this section.

Given this setup, a *necessary* condition for joint emigration is that the sum of the partners' net gains from migration to be positive. The Nash solution to this bargaining problem yields the following *sufficient* condition for joint emigration: the losses of any partner must not exceed household surplus h. Figure 1 illustrates the model predictions for all possible realizations of z_a and z_b. $A0$ is the region where each partner has private gains from migration, so that migration is a jointly desired by the couple. In case one partner, say a, has a migration incentive but b loses from

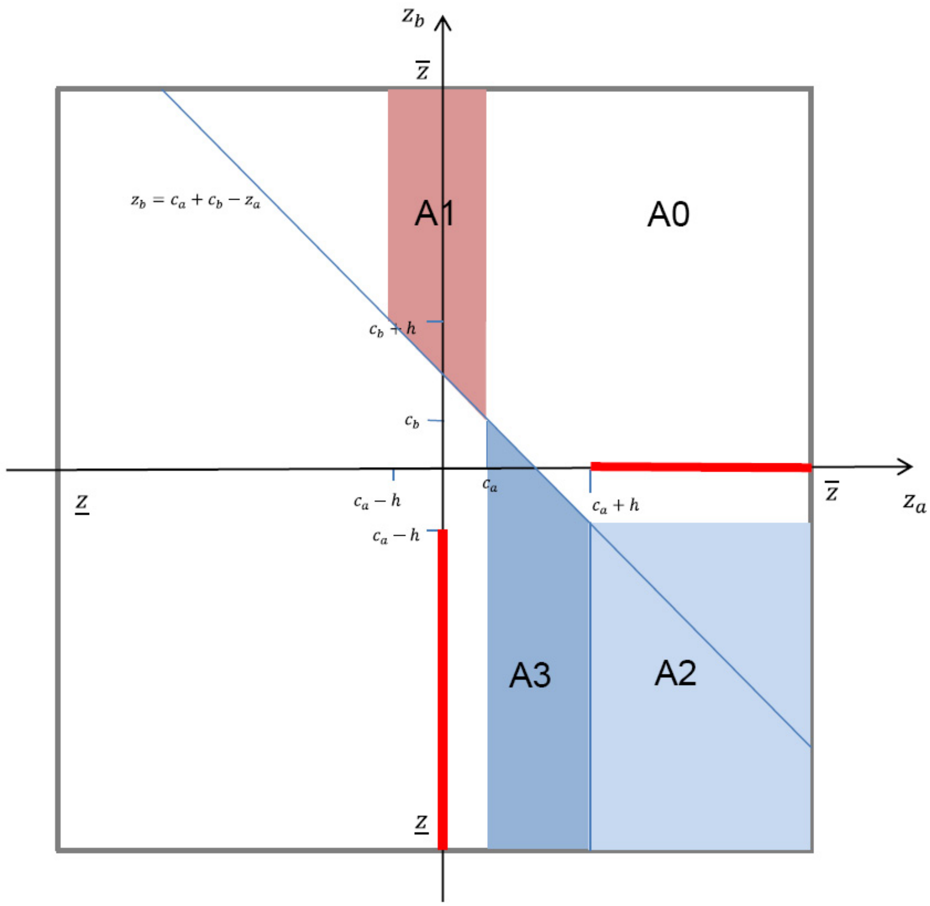

FIGURE 1. Possible realizations of migration gains z_a and z_b for both partners

migration, b becomes a tied mover, as long as b's losses are lower than household surplus h (Region $A1$). On the other hand, if b's losses from migration are too high, there is either a migration induced split up (Region $A2$), or both partners do not migrate and a becomes a tied stayer (Region $A3$).[9]

First stylized fact. The claim made by Eq.(2.7) is that the probability mass associated with region where the couple jointly migrates ($\mathbb{P}(A0) + \mathbb{P}(A1)$) is smaller than the probability associated to the region where, one partner migrates, for example $\mathbb{P}(A0) + \mathbb{P}(A2) + \mathbb{P}(A3)$, where we adopt the notation $\mathbb{P}(Ai) = \mathbb{P}((z_a, z_b) \in Ai)$. From this, it is clear that what needs to hold is that, given a level h, the probability of being a

[9]See Nikolka and Poutvaara (2013) for a formal derivation of these results

tied mover ($\mathbb{P}(A1)$) is smaller than of either being a tied stayer or of emigrating alone after split up ($\mathbb{P}(A2) + \mathbb{P}(A3)$). When is this condition likely to hold? If we assume that the joint distribution of gains from migration, z_a and z_b, is not too asymmetric and if migration costs are not too different, it will follow that the probability for joint emigration of a couple is lower than for a single person. Why do we expect it to be the case in Denmark? According to the United Nations Human Development Report 2011 Denmark is one of the most gender equal countries in the world. Income gaps between males and females are relatively small and we think that the symmetry assumptions are not too unrealistic in the case of couple migration from Denmark.

Second stylized fact. The usefulness of the second stylized fact which refers to the migration propensity of individuals from separated couples is a little more difficult to illustrate visually. It helps though to observe on Figure 1 the thick red lines on the right end of the z_a axis and on the low end of the z_b axis. They characterize the returns to migration of partners a and b, when a decides to migrate and b stays in the origin country. The insight here is that couples that experience separation because of migration incentives are the one with the most extreme returns to migration. After observing this, it is useful to condition on the event that one partner does not migrate after couple separation, say b; that is, in case of split up motivated by migration, we are sure that b is the partner with extreme negative returns and a the partner with extreme positive returns. Note now that in the model, couple separation is motivated by two reasons, either the household surplus is negative, or a migrates alone. Denote by H the random variable representing the household suplus of a couple, the event $Sep_a \equiv \{M_a = d \mid Y_{db=0}\}$ is the union of two disjoint events

$\{H < 0\} \cup \{(z_a, z_b) \in A2; H \geq 0\}$. It follows that[10]:

$$
\begin{aligned}
\mathbb{P}(Y_{da} = 1|M_a = d, Y_{db} = 0) &= \mathbb{P}(Y_{da} = 1|Sep_a) \\
&= \frac{\mathbb{P}\left(Y_{da} = 1; H < 0\right) + \mathbb{P}\left(Y_{da} = 1; A2; H \geq 0\right)}{\mathbb{P}(Sep_a)} \quad (2.9) \\
&= \frac{\mathbb{P}\left(Y_{da} = 1; H < 0\right) + \mathbb{P}\left(A2; H \geq 0\right)}{\mathbb{P}(Sep_a)} \quad (2.10) \\
&= \mathbb{P}\left(Y_{da} = 1|H < 0; Sep_a\right)\mathbb{P}(H < 0|Sep_a) \\
&\qquad + \frac{\mathbb{P}\left(A2; H \geq 0\right)}{\mathbb{P}(Sep_a)} \quad (2.11) \\
&= \mathbb{P}\left(Y_{da} = 1|H < 0; Sep_a\right)\mathbb{P}(H < 0|Sep_a) \\
&\qquad + 1 - \mathbb{P}\left(H < 0|Sep_a\right) \quad (2.12)
\end{aligned}
$$

Eq.(2.9) follows by application of the Bayes' rule and from the definition of the event Sep_a. Eq.(2.10) uses the fact that $\{(z_a, z_b) \in A2\}$ implies $\{Y_{da} = 1\}$. Eq.(2.11) applies a second time the Bayes rule to the first term on the right-hand side. Eq.(2.12) follows by definition of the event Sep_a.

Now, to retrieve the inequality of interest, it suffices to notice that:

$$
\begin{aligned}
\mathbb{P}(Y_{da} = \; 1 \; &|M_a = d, Y_{db} = 0) - F(Y_a = 1|M_a = c) \\
= \; &\left[\mathbb{P}\left(Y_{da} = 1|H < 0; Sep_a\right) - F(Y_a = 1|M_a = c)\right]\mathbb{P}\left(H < 0|Sep_a\right) \\
&+ \left(1 - F(Y_a = 1|M_a = c)\right)\left(1 - \mathbb{P}\left(H < 0|Sep_a\right)\right) \quad (2.13)
\end{aligned}
$$

Since we have here a convex combination of two quantities, $1 - F(Y_i = 1|M_i = c)$ being strictly positive, the sign of the difference will crucially depend on the sign of $\mathbb{P}\left(Y_{da} = 1|H_{ab} < 0; Sep_a\right) - F(Y_i = 1|M_i = c)$. If the latter is positive, that is, if split up has a positive effect on partners' mobility, as expected from the literature, the convex combination will be positive. Even if the difference is negative, the convex combination might still be positive if a sufficient proportion of couple separation is motivated by migration incentives[11].

To sum up the above analysis, we argue that if we consider the population of separated couples, we will find more couples with extreme returns in the population of

[10]For simplicity, we omit the conditioning on the X variable.

[11]This is because $1 - F(Y_i = 1|M_i = c)$ will be close to 1.

separated couples than in the rest of the couple population. Conditional on observing one member of the separated couple not migrating, it follows that migration should be observed more often among separated couples than in the rest of the couple population. The above argument requires, either that separated couple have somewhat similar or larger returns as the rest of the couple population, or that a sufficient proportion of separation is motivated by migration incentives.

The insights gained from this bargaining framework can be generalized to allow for cases where the commitment is not perfect in the couple. Indeed, the relaxation of the hypothesis of perfect commitment will lead, first, to lower couple migration propensity, so that Eq.(2.7) is even more likely to hold, and, second, to higher probability of couple separation (see Lundberg and Pollak, 2003). In this latter case, the migration incentives of the partner choosing couple separation migration will still be higher than those of the average individual, as long separation has a positive (or neutral) effect on international migration.

Note that the bounds rest on different types of conditions, the lower bound on the relative symmetry of couple members, and the upper bound on the positive association between mobility and couple separation. One way to assess the bounds' validity is to check whether they cross each other. In Table 2, we recompute the bounds using our new restrictions for the same population as in Table 1. Note that the upper bound is

Table 2. Bounds on $P(Y = 1|M = c)$ using the stylized facts and single migration rate

Educational attainment	LB	UB	Singles
Low education	0.04%	0.21%	0.19%
Lower middle educ.	0.07%	0.51%	0.42%
Upper middle educ.	0.12%	0.82%	0.45%
High education	0.28%	1.88%	1.11%

drastically reduced. Note also that the bounds do not cross, adding to the credibility of our assumptions. The bounds do not seem to be unreasonably wide either: for example, the bounds on the migration probability of the high educated individuals do not overlap with the one of the low educated. These patterns are consistent for different age group and both genders. *Prima facie*, we find no hint of selection. The observed single migration propensity belongs to the interval. Nevertheless, there is an additional payoff of using (separated or not separated) couples' migration behavior as bounds. We can exploit those characteristics of the couples that are not related to the private incentives of migration (as we would with an IV) or that monotonically move the bounds without violating the dominance (as with an MIV). The additional refinements are explored in the subsequent section.

2.3.4. *Additional Refinements.*

Lower bound. With regard to the lower bounds, it is tempting to assume that mating systematically decreases the migration probability of the partners, i.e.

$$F(Y_i = 1|M_i = c, x_i) \geq \max_{x_j} \mathbb{P}(Y_{ci} = 1|M_i = c, x_i, x_j) \quad \forall x_i \tag{2.14}$$

where we denote as Y_{ci} the decision of the couple to which i belongs and as X_m the observable characteristics of the husband. We will refer to Condition (2.14) as the *strong refinement on the lower bound*.

That mating decreases migration probability, irrespective of the partner characteristics is a rather controversial assumption. In the Appendix, we exhibit a special case in our model where Condition (2.14) will fail (Proposition 1 in the Appendix). Namely, we show that, as the pre-migration difference in income is large, a couple has higher likelihood of migration than the low-income earner. In other words, one partner might be found more mobile if she belongs to a very asymmetric couple than what she would actually be in the counterfactual case of no-marriage. Since we expect from the theory that this assumption will not hold, the strong refinement assumption could help us to gauge the quality of our bounds. If they are not too wide, they will cross under this assumption and reject it. Reassuringly, we find in Section 4 that the bounds derived from the strong refinement cross each other.

From the special case in Appendix, we can also prove, under fairly general conditions, that the probability of migration of a couple is always lower than the probability of migration of the high-income earner in the couple (Proposition 1). This result gives the intuition for the alternative, weaker restriction that we exploit:

$$F(Y_i = 1|M_i = c, x_i) \geq E_W \max_{W_m \leq W_i} \mathbb{P}(Y_{ci} = 1|M_i = c, x_i, x_m, W_i, W_m) \quad \forall x_i, x_m \tag{2.15}$$

where w_m stands for the earnings of the partner, and w_i, i's earnings. We refer to Condition (2.15) as the *weak refinement on the lower bound.*

The rationale here is that, in a couple, the high-income earner will be the one that is more often tied-stayed or migrate alone, than tied-moved by the partner. A sufficient condition for Eq.(2.15) to hold is that the high-income earner receives comparatively the same or the best job offers from abroad.

Upper bound. With regard to the upper bound, an equally tempting assumption is that irrespective of their partners' characteristics, individuals who decide to live in separate location in order to reap the benefits of migration have higher unobserved returns to migration than the average married individuals, i.e.

$$F(Y_i = 1|M_i = c, x_i) \leq \min_{x_j} \mathbb{P}(Y_{di} = 1|M_i = d, x_i, x_j, Y_{dj} = 0) \quad \forall x_i \tag{2.16}$$

We will refer to Condition (2.16) as the *strong refinement on the upper bound.* In our view, this hypothesis is less debatable than Condition (2.14), but we realize that it might be violated.

A more conservative approach is to consider variables directly influencing the decision to move after dissolution but not the private migration incentives. The time spent as a couple before separation appears a good candidate to satisfy this requirement[12].

$$F(Y_i = 1|M_i = c, x_i) \leq \min_{V_{ij}} \mathbb{P}(Y_{di} = 1|M_i = d, x_i, V_{ij}, Y_{dj} = 0) \quad \forall x_i \tag{2.17}$$

We will refer to Condition (2.17) as the *weak refinement on the upper bound.*

[12]A possible concern is that longer relationships might be a distinct characteristics for couples with high household surplus. These couples would be the most likely to migrate together, and the least likely to break up because of the migration decision. We can therefore make the alternative assumption that the migration propensity of the members of a separated couple is monotonically increasing in the length of the relationship, as with an MIV. Our findings are not changed under this alternative assumption.

3. Empirical analysis

3.1. **Description of Data.** We use data from the full Danish population register from 1990 to 2005. For each year in this period we restrict our attention to individuals aged 25 to 39 who are Danish citizens. In the subsequent analysis we consider three age-groups separately: 25-29, 30-34 and 35-39. Moreover we categorize individuals according to their educational attainment. We distinguish between low education (primary education), lower middle education (secondary education), higher middle education (medium cycle higher education) and high education (college degree). We exclude individuals for whom we have no information on educational attainment.

Our analysis compares for each cross-section year individuals living as singles with those living in opposite-sex couples. We use information on whether a person is cohabiting or married to distinguish between couples and singles. To be considered a couple, individuals have to be with the same partner at least for one year. We also analyze the group of separated partners, which are defined as individuals who hab been been living in a couple for at least one year but are not cohabiting with or married to the same partner in the following year.

We consider potential migration events in the year after a corresponding cross-section year. To be counted as an emigrant, a single must have left Denmark and stayed abroad for at least three years. Our analysis excludes emigrants to the autonomous Danish territories of the Faroe Islands and Greenland because the characteristics of these emigrants might systematically differ from migrants to other destination countries. We consider couples to have jointly emigrated if both partners leave Denmark for the same destination country within the same year and do not return for three years. Separated migrants are those who leave the country alone while their partner stays in Denmark either as a single or with a new partner.

We also observe the number of children younger than 18 living in a household, the relationship duration of couples as and the past relationship status of an individual's parents. We use income data to infer who is the primary earner in a couple.

3.2. **Descriptive Statistics.** Table 5 shows the educational distribution for females and males in the population. Couples differ from singles in terms of educational achievement. For example, males who are in a couple are relatively more often highly educated than male singles. This is also the case among females, apart from the high education category: Female singles have more often a higher education degree than females in couples.

Moreover, descriptive evidence suggests that couples and singles differ not only with respect to observable but also unobservable characteristics. We run a Mincer wage regression separately for males and females controlling for age, experience and experience squared. Figures 3 and 2 plot the distribution of the residuals of couples against those of singles, for males and females separately. We test whether the residuals for singles and couples are drawn from the same distribution using the Kolomogorow-Smirnov test for distributional equality. The test rejects distributional equality for both, females and males. Figures 3 and 2 show for positive earnings residuals that the distribution of singles dominates that of couple members, in particular for female individuals. A possible explanation is that jobs with high earnings require higher flexibility and mobility for a given education and experience experience level. Couple members might be more restricted than singles with respect to these unobserved characteristics. This might be reflected in their occupational choice as well as in career perspectives within a certain profession.

On the other hand, for negative earnings residuals the distribution of couples seems to dominate that of singles. This is, in particular, the case for males. We take this as suggestive evidence for selection into couples and into higher residual earnings based on similar underlying characteristics, for example attractiveness (see e.g. Lopez Boo, Rossi, and Urzua (2013) for evidence of a labor market premium due to attractiveness). We may also observe here the counter-effect of location ties of couple members discussed above: Individuals in couples who are tied to a location or select into certain occupations forgo earnings, on the one hand, but might also be more productive than their single peers, on the other hand, because of different specialization opportunities. Another potential effect of how family ties shape individual labor market outcomes might be incentives to under-invest in human capital. This could also result in higher unobserved ability after controlling for education and experience. Overall, the distributions of earnings residuals support our suggestion that singles and individuals in couples differ also in unobservable characteristics. Former literature has shown the role of these characteristics for individual migration decision (see Borjas (1987) for migration to the US and Kauppinen, Borjas, and Poutvaara (2013) for emigration from Denmark).

Figures 4 and 5 show migration rates of males and females by education groups. We present emigration rates for couple members, singles and separated partners. For all groups the likelihood of migration seems to increase with education. Migration rates for couples are persistently lower than for singles. Separated individuals, particularly men, have higher migration rates compared to those of singles. For example, the emigration rate for highly educated single males is 1.2%, for separated males it is almost 2.5%.

Table 6 sheds some more light on the role of observable characteristics in emigration, separately for single males, single females, and couples. We present estimation results from a Probit model explaining emigration of singles and couples as defined above with a binary dependent variable. As independent variables we include the age and education categories. For singles emigration is more likely among higher educated individuals, for both males and females. As suggested by Table 5 we also find that couple migration is positively associated with both partners' educational achievement: the higher educated are more likely to leave the country. However, higher male education seems to increase the likelihood of migration of the couple more than does female education. Additionally, we observe a higher likelihood of couple migration the younger the male partner. The findings are more ambiguous when it comes to the female partner's age. Among singles, migration probability decreases clearly with age for both males and females. In general, the results suggest that international migration by Danish couples is more responsive to male than to female age and education.

3.3. **Estimation Methodology.** The bounds we propose belong to the class of intersection bounds. For details on the inference methodology, the reader is referred to Chernozhukov, Lee, and Rosen (2013), henceforth CLR. CLR propose bias-corrected estimators of the upper and lower bounds, as well as confidence intervals[13]. Their approach employs a precision correction to the estimated bounding functions before applying the supremum and infimum operators. They achieve this by adjusting the estimated bounding functions for their precision by adding to each of them an appropriate critical value times their pointwise standard error.

We implement this procedure using the Stata code described in Chernozhukov, Kim, Lee, and Rosen (2013). All our tests use the parametric estimator. Note that because of the computational limitation, we draw a random sample from the population of non-migrants and adjust the parametric estimator with an appropriate weighting matrix. For simplicity, we compute $F(Y = 1|M = s, X = x)$ for each x from the whole population andl consider this probability mass to be a constant for each x.

3.3.1. *A Proposal for Constructing Bounds from an Unconditional MIV.* CLR propose a methodology for constructing intersection bounds in the case of the conditional MIV, that is, without the expectation operator. In our case, this proposal allows us to construct bounds such that:

$$F^l(Y = 1|M = s, X, Z = z) \leq F(Y = 1|X, Z = z) \leq F^u(Y = 1|M = s, X, Z = z) \tag{3.1}$$

for each z, and F^l and F^u are given respectively, by Eq. (2.4) and (2.5). These bounds are of limited use for our test of selection. Only in the case where $F(Y = 1|M = s, X, Z = z)$ fails to fall within these bounds, for all possible z on the support of Z, can we reject the hypothesis of no selection. The difficulty in constructing valid confidence regions for the unconditional MIV case has to do with the existence of the

[13]CLR note two reasons why estimation of and inference on intersection bounds is complicated: first, because the bound estimates are suprema and infima of parametric or nonparametric estimators, closed-form characterization of their asymptotic distributions is typically unavailable or difficult to establish. Second, since sample analogs of the bounds are the suprema and infima of estimated bounding functions, they have substantial finite sample bias, and estimated bounds tend to be much tighter than the population bounds.

min / max operator, within the conditional expectation. Note that Eq. (2.15) shares the same feature, a max operator within an expectation. In this particular case, a conditional version of the inequality is not available since we cannot observe the wage of a partner for a single individual.

We propose a two-stage procedure to retrieve valid confidence regions for the unconditional MIV case, in the event Z is a discrete random variable. The idea is to first apply a precision correction to the term within the min / max operator before proceeding with the usual CLR procedure with the expectation. In Appendix C, we provide a motivation for the procedure, and in Appendix D some simulation results to explore its coverage properties and to compare its performance with that of an inference procedure we describe as a "naïve" procedure. The results highlights two points of interest: (i) the naïve procedure exhibits some problematic size distortion, while the two-stage procedure provides valid confidence regions; and (ii) the proposed two-stage inference procedure provides conservative confidence regions (particularly with small samples). Nevertheless, the two-stage procedure might retain some information of the MIV, hence, tightening the bounds.

3.4. **Variance Decomposition: Implicit Bounds.** To assess each spouse's contribution to the family migration decision, we follow Tenn's (2010) methodology and perform a variance decomposition. Consider the following reduced form model:

$$Y_c = I(X_f\beta_f + X_m\beta_m + \varepsilon_c > 0) \text{ where } \varepsilon_c \sim N(0,1) \tag{3.2}$$

The above single index equation describes the couple migration decision. Eq. (3.3), describes each partner's migration decision in the event (s)he is not tied[14].

$$Y_g = I(X_g\tilde{\beta}_g + \varepsilon_g > 0) \text{ where } \varepsilon_g \sim N(0,1), \text{ for } g = f, m \tag{3.3}$$

[14]Note that in a departure from Tenn's framework, we consider that β_g differs from $\tilde{\beta}_g$ not only as a result of the normalization imposed on the variance of the unobservable characteristics. The above specification accounts for different possible decision mechanisms (unitary model, bargaining), including non-efficient ones; for example the possibility of a couple separation even if the sum of private surplus is positive.

To calculate the explanatory power of the female partner relative to male partner's, we compute the ratio:

$$r = \sqrt{\frac{Var\left(X_f \beta_f\right)}{Var\left(X_m \beta_m\right)}} \tag{3.4}$$

A value $r = 1$ implies that both partners have equal explanatory power in the family migration decision. In our population, we find $r = 0.34$, suggesting a weak dominance of the female partner's observable characteristics in the migration decision.

However, the above ratio does not control for potential differences in migration preferences. As noted by Tenn, the low explanatory power of female characteristics could be the result of a gender asymmetry in migration preference rather than illustrative of the aggregation rule of individual preferences within the couple. To address this issue, Tenn proposes to correct the above ratio with a counterfactual variance decomposition, i.e. to compute the ratio:

$$\tilde{r} = r \div \sqrt{\frac{Var\left(X_f \tilde{\beta}_f\right)}{Var\left(X_m \tilde{\beta}_m\right)}} \tag{3.5}$$

The denominator here captures the asymmetry in private incentive for migration, that is the asymmetry in the decision individuals would have taken had they been single. This private incentive is unobserved,however. To circumvent the partial observability problem, Tenn assumes that the agents in couples would have behaved similarly to their single counterparts had they not been in a relationship.

As discussed in the previous section, selection bias implies that singles are poor proxies for individuals in couples. We thus propose to derive "implicit bounds" on the ratio of variances of the mean return to migration, from the bounds motivated in Section 2.3.

Recall the bounds advised in Section 2.3 on the quantity $F\left(Y = 1|M = c, x\right)$:

$$\begin{aligned} F\left(Y_i = 1|M_i = c, x\right) &\leq \max_{X_j} \mathbb{P}\left(Y_{ci} = 1|M_i = c, x, X_j\right) \equiv LB(x) \\ F\left(Y_i = 1|M_i = c, x\right) &\leq \min_{X_j} \mathbb{P}\left(Y_{di} = 1|M_i = d, x, X_j, Y_{dj} = 0\right) \equiv UB(x) \end{aligned}$$

for all x on the support of X. From the structural model 3.2, we can rewrite the above equation:

$$LB(x) \leq \Phi\left(x\tilde{\beta}_g\right) \leq UB(x) \tag{3.6}$$

where $\Phi(v)$ is the cumulative distribution function of a random variable with standard normal distribution. It follows that the parameter $\tilde{\beta}_g$ is only partially identified, that is the set of values that could generate the observed data generating process might not be reduced to a singleton. Denote as Θ the set of all values satisfying Eq. (3.6). From the sample, we can then compute the following two quantities

$$\begin{aligned} LBV_g &= \min_{\tilde{\beta}_g \in \Theta} Var\left(X_g\tilde{\beta}_g\right) \\ UBV_g &= \max_{\tilde{\beta}_g \in \Theta} Var\left(X_g\tilde{\beta}_g\right) \end{aligned}$$

The denominator in Eq. (3.5) is then bounded as follows:

$$\frac{LBV_f}{UBV_m} \leq \sqrt{\frac{Var\left(X_f\tilde{\beta}_f\right)}{Var\left(X_m\tilde{\beta}_m\right)}} \leq \frac{UBV_f}{LBV_m} \tag{3.7}$$

Note that if 1 belongs to this interval, we cannot reject gender symmetry in private migration incentives. The bounds on the ratio of interest $\tilde{r}$ follow trivially.

We now turn to the results of our test procedure.

4. Results

4.1. **Test of Selection.** To minimize the number of tests that need to be conducted, we first successively test:

$$H_0^l : \qquad \max_x \left(LB(x) - F(Y=1|M=s,x)\right) \leq 0 \text{ and} \tag{4.1}$$

$$H_0^u : \qquad \min_x \left(UB(x) - F(Y=1|M=s,x)\right) \geq 0 \tag{4.2}$$

where $LB(x)$ and $UB(x)$ summarize, respectively, the lower bound and the upper bound derived in Section 2 under alternative assumptions. If either H_0^l or H_0^u is rejected, say for example under the IV assumption, it means that there is at least one age-education cell where singles and couple members differ on unobservable characteristics (assuming that the IV-assumption is correct).

TABLE 3. One-sided CI ends for $\max_x (LB(x) - F(Y=1|M=s,x))$ and $\min_x (UB(x) - F(Y=1|M=s,x))$ respectively, for males and females under alternative assumptions.

	Males		Females	
one-sided CI level	$\max LB - F$	$\min UB - F$	$\max LB - F$	$\min UB - F$
IV				
90% CI	-0.0007	0.2812	0.0001	0.5427
95% CI	-0.0007	0.2861	0.0001	0.547
99% CI	-0.0007	0.2954	0.0000	0.5552
Economically motivated + IV				
90% CI	-0.0006	0.0008	0.0002	0.0004
95% CI	-0.0007	0.0009	0.0002	0.0005
99% CI	-0.0007	0.0012	0.0001	0.0007
MIV				
90% CI	-0.0014	0.2263	0.0001	0.2514
95% CI	-0.0014	0.2263	0.0001	0.2514
99% CI	-0.0014	0.2263	0.0001	0.2514
Economically motivated + MIV				
90% CI	-0.0013	-0.0003	0.0003	-0.0017
95% CI	-0.0013	-0.0003	0.0003	-0.0016
99% CI	-0.0013	-0.0003	0.0003	-0.0016
Economically motivated + Strong refinement				
90% CI	-0.0006	0.0023	0.0025	-0.0007
95% CI	-0.0007	0.0026	0.0024	-0.0005
99% CI	-0.0007	0.0031	0.0023	-0.0000
Economically motivated + Weak refinement				
90% CI	-0.0007	0.0044	0.0020	0.0006
95% CI	-0.0007	0.0046	0.0020	0.0008
99% CI	-0.0008	0.0051	0.0020	0.0013

We perform the tests separately for males and females. Table 3 summarizes the results for each gender group. Columns (1) and (3) ((2) and (4)) show the lower (upper) end of a one-sided confidence region on the quantity in Eq. (4.1) (Eq. (4.2)), under alternative assumptions. A rejection of the null H_0^l (H_0^u) is equivalent to this lower (upper) end being above 0 (below 0). Note that if the two ends cross, say, for example, for the "strong refinement" assumption, it is important to check whether this happens within the same age-education cell. Such crossing is indicative that the hypothesis, in our example the "strong refinement"-assumption, is incorrect. This is indeed the case, as expected, for the "strong refinement" assumption, in the female population. Entering a couple might increase the migration probability of one partner.

Consider first the female population. H_0^l is rejected for all specifications. This result means that there is at least one age-education cell where couple members have higher private migration incentives than singles. That the rejection occurs under very different assumptions makes us confident of the robustness of the result. According to the previous discussion in Section 2 of the potential selection mechanisms, our findings suggest that either the specialization advantage or the positive selection on unobserved characteristics valued on both the marriage and foreign labor markets is at play. To disentangle the two, we can look at the age-education cell(s) driving the rejection. After repeating the testing procedure for each age-education cell, we find that married low educated and older women (aged between 35-39) show migration incentives significantly different from those of their single counterparts. We doubt that this group os high in specialization and we also suspect that those who would be found less attractive on the marriage market are overrepresented in this subgroup.

With respect to H_0^u, the null is rejected in two cases: the economically motivated upper bound coupled either with the MIV or with the strong refinement assumption. Note, however, how close the bound is from zero in the case where we combine the economically motivated bound with the IV. We will return to this finding below. Assuming that Stylized fact 2 is valid, we can conclude that there is at least one age-education cell in which couple members have lower private migration incentives than singles. Again, we investigate which subgroup is driving the rejection of the null. Repeating the testing procedure for each age-education cell, we find that the youngest married women (aged between 25-29) have migration incentives significantly lower than from those of their single counterparts. This finding confirms our idea that migration-prone individuals are sorted out of marriage.

Consider next the male population. The only case of rejection is related to the upper bound, which is smaller than the migration propensity of singles under the "economically motivated + MIV" assumption. Therefore, the test procedure provides only scant evidence of selection in the male population. Note again, how close the upper end of the confidence region is to 0 when we consider the combination of the economically motivated bound with each remaining assumption[15]. In fact, a test for $\min_x \left(UB(x) - F(Y=1|M=s,x)\right) = 0$ will not be rejected under these assumptions. To interpret this result, recall that the upper bound is relative to the observed migration propensity of members of separated couples. As discussed in Section 2.3, individuals with extreme migration returns are overrepresented in this subpopulation. Separated individuals thus could be viewed as a selected subpopulation among the individuals in a couple in that their migration propensity *strictly* dominates that of the partners who remained in co-location. The results of Table 3 suggest that there is at least one age-education cell where singles have the same migration propensity as these separated individuals. In other words, certain singles also select themselves out of marriage due to private migration incentives. In retrospect, it is not surprising that the private migration propensities of singles and the separated coincide since both groups have self-selected out of marriage.

As a robustness check, we include among the controlled-for individual characteristics, the labor income of each agent and repeat our sequence of tests. Controlling for this additional characteristic helps to indirectly account for any individual specific productivity that might drive part of the selection. Kauppinen, Borjas, and Poutvaara (2013) show that international migrants from Denmark are positively selected on their individual-specific productivity, as measured by the residual of the Mincerian regression. For computational reasons, we limit ourselves to the annual labor income quintile to which the individual belongs to. The inclusion of this variable does not significantly alter our conclusion. In particular, in the female population, we still reject the null H_0^l in almost all specifications. However, the results are less supportive

[15]For some of the samples we draw in the population, this upper bound is even negative.

of the sorting effect. That the sorting effect is less apparent, once labor income is controlled for, suggests that singles with high individual specific-productivity (and thus with the highest returns to migration) delay couple formation so as to reap the benefits of the migration opportunity.

Overall, our data support the idea that singles' private migration incentives are not a good proxy for thos of their married counterparts. Moreover, the underlying selection mechanisms differ by gender, in violation of the assumption required for using Tenn's (2010) procedure. Could this selection bias reverse the result we find when using singles as the counterfactual in the variance decomposition exercise?

4.2. **Bounds on Variance Decomposition of Couples' Migration Decision.** We now repeat Tenn's (2010) procedure, that is, compute Eq.(3.5), but replace the denominator by the ratio of the variance of the mean returns of singles. In our population, the numerator r is estimated at 0.3467. Since r measures the explanatory power of the male partner's characteristics relative to the explanatory power of the female partner's characteristics, our estimates indicate that the female's characteristics have far less explanatory power than the male's. The denominator in Eq. (3.5) controls for potential gender differences in the migration preferences of men and women. Computed for singles, it yields an estimated value of 1.2046. Interestingly, these estimates suggest that the gender asymmetry is such that the observed characteristics of single females are more important to their migration decisions than the characteristics of single males are to their migration decisions. Indeed, in some age-education cell, single females exhibit a significantly higher migration propensity than do males. The estimated value for $\tilde{r}$ is then 0.2878.

We now compute the bounds on the ratio of the variance of the mean return to migration, as established in Eq. (3.7) in Section 3.4. Throughout this section, we use the 95% two-sided confidence region for the interval $[LB(x); UB(x)]$[16]. To construct Θ from Eq. (3.6), we generate a grid on the parameter space $\mathbb{R}^6$ with approximately 5×10^6 nodes. The grid is refined around the points $(A_x^T A_x)^{-1} A_x^T \Phi^{-1}(LB(A_x))$ and $(A_x^T A_x)^{-1} A_x^T \Phi^{-1}(UB(A_x))$, where A_x is the matrix of all possible combinations of the independent variable. For each point in Θ, we compute for the full population

[16]Using the median-unbiased estimator does not change the main findings.

of married individuals the variance of the mean return to migration, separately for males and females. The quantity LBV_g (UBV_g) is then the minimum (maximum) over the points in Θ.

Table 4. Bounds on the variance of mean return to migration and their ratio by gender

	Econ.	Econ. + MIV	Econ. +weak ref	Econ. +strong ref.	All except strong ref.
(A) Variance of mean return to migration					
Female					
Lower bound LBV_f	0.0015	0.0007	0.0007	0.0180	0.0028
Upper bound UBV_f	0.1242	0.1465	0.0586	0.0336	0.0491
Male					
Lower bound LBV_m	0.0005	0.0014	0.0005	0.0005	0.0014
Upper bound UBV_m	0.165	0.165	0.2412	0.1576	0.190
(B) Ratio female to male					
Lower bound $\sqrt{LBV_f/UBV_m}$	0.0952	0.065	0.0534	0.3383	0.1209
Upper bound $\sqrt{UBV_f/LBV_m}$	15.53	10.067	10.667	8.0836	5.8282
(C) Ratio female to male - Using single male as counterfactual					
Lower bound $\sqrt{LBV_f/V_{sm}}$	0.151	0.151	0.1771	0.7681	0.3014
Upper bound $\sqrt{UBV_f/V_{sm}}$	2.1887	1.3843	2.0153	1.0489	1.2672

Table 4 summarizes the results of this procedure, using alternatively the different assumptions we entertained in the previous section (economically motivated, weak refinement and MIV), as well as a combination of all three of them[17]. Note that for the assumption "economic + strong refinement" (Column 4), the age-education cell where the bounds cross is replaced by its counterpart without the strong refinement.

[17]As we argued in Section 2, the validity of the IV assumption is disputable and we thus exclude it from the present analysis. For the same reason, we treat the strong refinement assumption separately.

Note also that, since the CLR procedure is not required in the case where we construct the economic bounds, we use the full population of observations. Finally, to mediate the effect of the over-correction from the TSCLR on the lower bound using the MIV assumption, we compute the lower bound without considering the MIV assumption.

Panel A of Table 4 displays by gender the variance of the mean return to migration. The magnitude of these variances is subject to relatively large variations as one moves through the identified set. As a result, the ratio computed in Panel B ranges from about one-tenth to a hundred times this value. Arguably, this is more a shortcoming of the implicit bounds than an effect of selection. The bounds are too wide to be informative. Since the bulk of the evidence for selection is found in the female population, and very little on the male population, a reasonable solution is to use single males as representative of the married males. The bounds on the ratio come from the bounds on the variance derived on females' return to migration. The results displayed in Panel C show a substantial reduction of the magnitude of the bounds, especially when considering the economically motivated bounds and their strong refinement (Column 4) and the combination of the three assumptions (Column 5). The equality of the private migration incentive cannot be rejected as the bounds always contain 1. However, the strong refinement assumption reveals a tilt toward values lower than unity, in contrast with the result obtained by comparing only singles' return to migration.

We now have all the elements necessary for assessing the relative importance of partners' characteristics in a couple's migration decision. From the outcomes of the above calculation, we can bound the quantity $\tilde{r}$ in Eq.(3.5). Again, using the singles as proxy for the married yields an estimated value of $\tilde{r} = 0.2878$. This result unambiguously leads us to conclude like Tenn that female partners private return to migration is a weak determinant of family migration.

Using instead the bounds leads to a somewhat different conclusion. Consider the bounds we derive from the combination of the economically motivated, weak refinement and MIV assumptions. We find $\tilde{r}$ to range between 0.2736 and 1.1503. In this case, we cannot reject the hypothesis that each partner's return to migration is equally weighted in the decision-making process. Only under the strong refinement assumption, do the bounds fall under 1, ranging between 0.3305 and 0.4514.

5. Conclusion

Identification of the individuals' private incentives is an important step toward understanding a family's decision-making process. Doing so, however, requires resolving issues related to the selection into couples and the inherent partial observability of each member's private incentive to migrate. To overcome this challenge, this paper introduced and motivated bounds on the private unobserved migration propensity of married individuals, thus making several important contributions to the literature on couple migration.

First, we test whether married individuals differ from single individuals in their migration propensity even after controlling for their observable characteristics (in our case age and education). The results of this test answers the question of whether singles' migration propensity provides a reliable counterfactual for their married counterparts unobserved migration propensity. Performing the test on a population of Danish citizen aged between 25 to 39 leads to the conclusion that married individuals are selected on unobservable characteristics, especially females. We argue that there are at least two mechanisms at play in this phenomenon. On the one hand, there are unobserved traits that are both of importance for the international mobility of an individual and the formation of a couple. On the other hand, those with a high return to international migration will delay couple formation.

Second, we assess the relative importance of male and female partners' characteristics in the decision to migrate. Our findings make clear that at least for the Danish population that we study, using singles as proxies for the married could be misleading. Unless one makes the arguably doubtful assumption that couple formation unambiguously decreases the female partner's migration propensity, one cannot reject the hypothesis that both partners' observed characteristics are equally weighted in the migration decision.

As a final contribution, we proposed a two stage procedure for retrieving valid confidence regions for bounds derived from unconditional MIV assumption, when the monotone instrument has a finite support. The two-stage inference procedure shows interesting finite sample properties in the Monte-Carlo simulations. Moreover, it was particularly useful in our test procedure.

References

Amato, P. R. (1996): "Explaining the intergenerational transmission of divorce," *Journal of Marriage and the Family*, pp. 628–640.

——— (2000): "The consequences of divorce for adults and children," *Journal of marriage and family*, 62(4), 1269–1287.

Angrist, J., E. Bettinger, and M. Kremer (2004): "Long-term consequences of secondary school vouchers: Evidence from administrative records in Colombia," Discussion paper, National Bureau of Economic Research.

Åström, J., and O. Westerlund (2009): "Sex and migration: Who is the tied mover?," unpublished manuscript.

Axelsson, R., and O. Westerlund (1998): "A panel study of migration, self-selection and household real income," *Journal of Population Economics*, 11(1), 113–126.

Blundell, R., A. Gosling, H. Ichimura, and C. Meghir (2007): "Changes in the distribution of male and female wages accounting for employment composition using bounds," *Econometrica*, 75(2), 323–363.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," *American Economic Review*, 77(4), 531–53.

Boyle, P., T. Cooke, K. Halfacree, and D. Smith (1999): "Gender inequality in employment status following family migration in GB and the US: The effect of relative occupational status," *International Journal of Sociology and Social Policy*, 19(9/10/11), 109–143.

Cherlin, A. J., K. E. Kiernan, and P. L. Chase-Lansdale (1995): "Parental divorce in childhood and demographic outcomes in young adulthood," *Demography*, 32(3), 299–318.

Chernozhukov, V., W. Kim, S. Lee, and A. Rosen (2013): "Implementing intersection bounds in Stata," Discussion paper, cemmap working paper, Centre for Microdata Methods and Practice.

Chernozhukov, V., S. Lee, and A. M. Rosen (2013): "Intersection bounds: Estimation and inference," *Econometrica*, 81(2), 667–737.

Costa, D. L., and M. E. Kahn (2000): "Power couples: Changes in the locational choice of the college educated, 1940–1990," *The Quarterly Journal of Economics*, 115(4), 1287–1315.

Duncan, R. P., and C. C. Perrucci (1976): "Dual occupation families and migration," *American Sociological Review*, pp. 252–261.

Feijten, P., and M. Van Ham (2007): "Residential mobility and migration of the divorced and separated.," *Demographic Research*, 17.

Feijten, P., and M. van Ham (2013): "The consequences of divorce and splitting up for spatial mobility in the UK," *Comparative Population Studies*, 38(2).

Frank, R. H. (1978): "Why women earn less: The theory and estimation of differential overqualification," *American Economic Review*, 68(3), 360–373.

Galichon, A., and M. Henry (2013): "Dilation bootstrap," *Journal of Econometrics*, 177(1), 109–115.

Gautier, P. A., M. Svarer, and C. N. Teulings (2010): "Marriage and the city: Search frictions and sorting of singles," *Journal of Urban Economics*, 67(2), 206–218.

Gemici, A. (2007): "Family migration and labor market outcomes," Ph.D. thesis, University of Pennsylvania.

Gemici, A., and S. Laufer (2011): "Marriage and cohabitation," unpublished manuscript.

Gram-Hanssen, K., and C. Bech-Danielsen (2008): "Home dissolution: What happens after separation?," *Housing Studies*, 23(3), 507–522.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2012): "Technology and the changing family: A unified model of marriage, divorce, educational attainment and married female labor-force participation," Discussion paper, National Bureau of Economic Research.

Greenwood, M. J. (1997): "Internal migration in developed countries," .

Grundy, E. (1985): "Divorce, widowhood, remarriage and geographic mobility among women," *Journal of Biosocial Science*, 17(4), 415–35.

Guler, B., F. Guvenen, and G. L. Violante (2012): "Joint-search theory: New opportunities and new frictions," *Journal of Monetary Economics*, 59(4), 352–369.

Henry, M., R. Méango, and M. Queyranne (2011): "Combinatorial bootstrap inference in partially identified incomplete structural models," Unpublished manuscript.

Jacobsen, J. P., and L. M. Levin (1997): "Marriage and migration: Comparing gains and losses from migration for couples and singles," *Social Science Quarterly*, 78(3), 688–709.

Kauppinen, I., G. Borjas, and P. Poutvaara (2013): "Self-selection of emigrants from a welfare state," unpublished manuscript.

Lopez Boo, F., M. A. Rossi, and S. S. Urzua (2013): "The labor market return to an attractive face: Evidence from a field experiment," *Economics Letters*, 118(1), 170–172.

Lundberg, S., and R. A. Pollak (2003): "Efficiency in marriage," *Review of Economics of the Household*, 1(3), 153–167.

Manski, C. F. (1995): *Identification problems in the social sciences*. Harvard University Press.

Manski, C. F., and J. V. Pepper (2000): "Monotone instrumental variables: With an application to the returns to schooling," *Econometrica*, 68(4), 997–1010.

Manski, C. F., G. D. Sandefur, S. McLanahan, and D. Powers (1992): "Alternative estimates of the effect of family structure during adolescence on high school graduation," *Journal of the American Statistical Association*, 87(417), 25–37.

Méango, R. (2014): "International student migration: A partial identification analysis," Discussion paper, CESifo Group Munich.

Mincer, J. (1978): "Family migration decisions," *Journal of Political Economy*, 86, 749–73.

Munk, M. D., M. Junge, and P. Poutvaara (2013): "International Migration of Power Couples," unpublished manuscript.

Nivalainen, S. (2004): "Determinants of family migration: Short moves vs. long moves," *Journal of Population Economics*, 17(1), 157–175.

Poirier, D. J. (1980): "Partial observability in bivariate probit models," *Journal of Econometrics*, 12(2), 209–217.

Rabe, B. (2011): "Dual-earner migration. Earnings gains, employment and self-selection," *Journal of Population Economics*, 24(2), 477–497.

Ribar, D. C. (2004): "What do social scientists know about the benefits of marriage?: A review of quantitative methodologies," (No. 998), IZA Discussion paper series.

Sandell, S. H. (1977): "Women and the economics of family migration," *Review of Economics and Statistics*, 59(4), 406–414.

Sorenson, O., and M. S. Dahl (2012): "Geography, joint choices and the reproduction of gender inequality," Unpublished manuscript.

Symon, P. (1990): "Marital breakdown, gender and home ownership: The owner-occupied home in separation and divorce," *Housing and Divorce*, pp. 110–138.

Tamer, E. (2010): "Partial identification in econometrics," *Annual Review of Economics*, 2(1), 167–195.

Tano, S., O. Westerlund, R. Nakosteen, and M. Zimmer (2014): "Back to the future: Migration, matching and the power couple phenomenon in Sweden," Discussion paper, Umeå University, Department of Economics, Unpublished manuscript.

Taylor, M. P. (2007): "Tied migration and subsequent employment: Evidence from couples in Britain," *Oxford Bulletin of Economics and Statistics*, 69(6), 795–818.

Tenn, S. (2010): "The relative importance of the husbands and wifes characteristics in family migration, 1960–2000," *Journal of Population Economics*, 23(4), 1319–1337.

Waite, L., and M. Gallagher (2002): *The case for marriage: Why married people are happier, healthier and better off financially*. Random House LLC.

Appendix A. Assessing the refinements

We exhibit a special case in our model in Section 2.3.3 where condition (2.14) will fail. The analysis also gives intuition for the weaker restriction (2.15). Note that normality of the distribution assumed below is not required for the following argument to be valid. In general, the class of absolutely continuous symmetric distributions will lead to Proposition 1. We assume that individual utility is linear and separable in wage and migration costs. Assume that z_i is the return to migration of an individual i with wage w_i

$$z_i = x_i w_i \tag{A.1}$$

where x_i is a random variable normally distributed with mean μ_i and standard error σ_i. We assume that $\mu_i < 0$, so that even with negligible costs of migration, half the population would not migrate. As previously, let c_i denote the cost of migration for individual i. i migrates if $z_i > c_i$. Assume that the random variables related to the members of the couple are such that

$$(x_a, x_b) \sim \mathcal{N}\left((\mu_a, \mu_b)\,; \Sigma\right) \text{ where } \Sigma = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix}. \tag{A.2}$$

σ_{ab} measures the covariance between the two variables. We denote as ρ their correlation, which measures the similarity of the opportunities for income growth from migration. If we think of endogamy, we expect ρ to be positive. We still assume that $\mu_a < 0$ and $\mu_b < 0$, so that even with negligible costs of migration, half the population would not migrate. We will denote by ν the random variable $x_a w_a + x_b w_b$ with mean μ_ν and variance σ_ν^2.

Assume for simplicity infinitely large household surplus. The couple's migration decision is then guided by the maximization of its expected lifetime return. They migrate when: $z_a + z_b = x_a w_a + x_b w_b > c_a + c_b$.

We show an intuitive result that when facing the same costs, and having returns on migration drawn from the same distribution, the probability of migration of a couple is lower than the probability of migration of the higher-income earner in the couple. Further, when the difference in salary is relatively small, the probability of migration of the low-income earner also exceeds the migration probability of the couple. On the other hand, when the difference in salary becomes relatively large, the couple has higher likelihood of migration than the lower-income earner. We later discuss possible relaxations of the assumptions that make the results more general.

Assumption 1 (Same distribution)**.** *Assume the following:*

(1) $c_a = c_b = c > 0$,
(2) $\mu_a = \mu_b = \mu < 0$,
(3) $\sigma_a = \sigma_b = \sigma$,
(4) $\sigma_{ab} = 0$.

Recall that we have $w_a \geq w_b$, that is a has a higher salary in the source-country than b. Appealing to some lemma, we show the following proposition.

Proposition 1. *Denote as P_{ab} the migration probability of the couple, and as P_i the migration probability of individual $i \in \{a, b\}$. Under Assumption 1:*

(i) $P_{ab} \leq P_a$
(ii) There exist $\delta > 0$, such that if $w_a/w_b \leq \delta$, then $P_{ab} \leq P_b$, else $P_{ab} > P_b$.

A.1. **Proof of Proposition 1.** The first lemma shows that the difference $P_{ab} - P_a$ is the difference between the set of realizations of return to migration for which a is a tied-mover and a is a tied-stayer.

Lemma 1. *Denote:*

$$\begin{aligned}\zeta_{ab}^{a-} &= \left\{x = (x_a, x_b) \in \mathbb{R}^2 : x_a w_a + x_b w_b - 2c \geq 0 \geq x_a w_a - c\right\} \\ \zeta_{a}^{ab-} &= \left\{x = (x_a, x_b) \in \mathbb{R}^2 : x_a w_a + x_b w_b - 2c \leq 0 \leq x_a w_a - c\right\}\end{aligned}$$

$$P_{ab} - P_a = \mathbb{P}\left(\zeta_{ab}^{a-}\right) - \mathbb{P}\left(\zeta_{a}^{ab-}\right) \tag{A.3}$$

Proof. It suffices to see that:

$$\begin{aligned}P_{ab} &= \mathbb{P}\left(x_a w_a + x_b w_b - 2c \geq 0; x_a w_a - c > 0\right) + \mathbb{P}\left(x_a w_a + x_b w_b - 2c \geq 0; x_a w_a - c \leq 0\right) \\ P_a &= \mathbb{P}\left(x_a w_a + x_b w_b - 2c \leq 0; x_a w_a - c \geq 0\right) + \mathbb{P}\left(x_a w_a + x_b w_b - 2c > 0; x_a w_a - c \geq 0\right)\end{aligned}$$

□

The next lemma states that for each point $x^{a-} \in \zeta_{ab}^{a-}$, we can associate a point $x^{ab-} \in \zeta_{a}^{ab-}$, such that the probability mass associated with x^{a-} is always lower than the probability mass associated with x^{ab-}

Lemma 2. *Let $x^{a-} \in \zeta_{ab}^{a-}$. There exists a bijection ψ from defined on ζ_{ab}^{a-} to ζ_{a}^{ab-}, such that $x^{ab-} \equiv \psi(x^{a-})$ and*

$$\phi(x^{a-}) \leq \phi(x^{ab-}) \tag{A.4}$$

Proof. We show that the symmetry with respect to a given hyperplane in $\mathbb{R}^2$ is the appropriate bijection ψ to consider.

Consider the plan generated by the vector (x_a, x_b) (i.e $\mathbb{R}^2$). As of yet, we still assume that $x_a \perp x_b$.

- Call (D_a), the horizontal line such that $x_a w_a - c = 0$.
- Call (D_{ab}), the tilted line such that $x_a w_a + x_b w_b - 2c = 0$.

These straight lines delimit the subset ζ_{ab}^{a-} and ζ_{a}^{ab-} and intersects each other at the point $\left(\frac{c}{w_a}, \frac{c}{w_b}\right)$. Most importantly, there exists a hyperplane $(D^{\perp})$ in $\mathbb{R}^2$ (a straight line), such that the two subsets are symmetric relative to $(D^{\perp})$. Note that:

$$\left(\frac{c}{w_a}, \frac{c}{w_b}\right) \in (D^{\perp})$$

$$\left(\frac{c}{w_a}, \frac{c}{w_b}\right) \succ_{R^2} \left(\frac{c}{w_a}, \frac{c}{w_a}\right)$$

In addition, let the following equation characterize $(D^{\perp})$.

$$f(x_a, x_b, w_a, w_b, c) = x_b - \beta_0 - \beta_1 x_a = 0$$

It is easy to check that the assumption $w_a \geq w_b$ implies that $0 \leq \beta_1 \leq 1$. Finally, define the lower contour set of level 0 the function f as:

$$\underline{D}_f = \left\{x = (x_a, x_b) \in R^2 : f\left(x, w_a, w_b, c\right) \leq 0\right\}$$

Let $x^{a-} \in \zeta_{ab}^{a-}$, and define x^{ab-} the point in ζ_{a}^{ab-} such that x^{a-} and x^{ab-} are symmetric with respect to $D^{\perp}$. To prove Eq. (A.4), it suffices now to show that :

$$\left\|x^{a-} - (\mu, \mu)\right\|_2 \geq \left\|x^{ab-} - (\mu, \mu)\right\|_2 \tag{A.5}$$

where$\|.\|_2$ is the usual Euclidian distance. This follows readily by noting that $(\mu, \mu) \in \underline{D}_f$. □

Proof of Proposition 1.i. By Lemma 1, it suffices to show that:

$$\mathbb{P}\left(\zeta_{ab}^{a-}\right) - \mathbb{P}\left(\zeta_{a}^{ab-}\right) < 0$$

We have:

$$\begin{aligned}
\mathbb{P}\left(\zeta_{ab}^{a-}\right) &= \int I\left(x^{a-} \in \zeta_{ab}^{a-}\right) d\Phi\left(x^{a-}\right) \\
&\leq \int I\left(x^{ab-} \in \zeta_{a}^{ab-}\right) d\Phi\left(x^{ab-}\right) \\
&= \mathbb{P}\left(\zeta_{a}^{ab-}\right)
\end{aligned}$$

The inequality follows by Lemma 2 and by the property of the bijection. □

The main point of the proof is that each point in ζ_{ab}^{a-} corresponds to a point in ζ_a^{ab-} closer to the mean w.r.t. to a well-defined distance. Assumption 1 can be relaxed in several significant directions.

The proof of Proposition 1.ii follows with analogous reasoning. The main insight is that for a relatively small difference in salaries, the mean of the bivariate distribution will belong to the upper contour set of a given function characterizing a plan of symmetry and when the difference is relatively large, to the lower contour set of the same function. Consider:

$$\begin{aligned} \zeta_{ab}^{b-} &= \left\{x = (x_a, x_b) \in \mathbb{R}^2 : x_a w_a + x_b w_b - 2c \geq 0 \geq x_b w_b - c\right\} \\ \zeta_{b}^{ab-} &= \left\{x = (x_a, x_b) \in \mathbb{R}^2 : x_a w_a + x_b w_b - 2c \leq 0 \leq x_b w_b - c\right\} \end{aligned}$$

Call (D_b), the vertical line such that $x_b w_b - c = 0$. These straight lines delimit the subset ζ_{ab}^{b-} and ζ_b^{ab-} and intersect each other at the point $\left(\frac{c}{w_a}, \frac{c}{w_b}\right)$. Again, there exists an hyperplane $(D_b^{\perp})$ in $\mathbb{R}^2$ (a straight line), such that the two subsets are symmetric relative to $(D_b^{\perp})$. The following equation characterizes the $(D_b^{\perp})$.

$$h(x_a, x_b, w_a, w_b, c) = x_b - \gamma_0 - \gamma_1 x_a = 0$$

It is easy to check that the assumption $w_a \geq w_b$ implies that $\gamma_1 > 1$. Finally, call $\underline{D}_h(w_a, w_b)$ the lower contour set of level 0 the function h and $\overline{D}_h(w_a, w_b)$ the upper contour set of level 0 the function h. Here we emphasize the dependence of the contour sets on (w_a, w_b).

The full argument is tedious and unnecessary. The easiest way to understand the result is to consider the two extreme cases. Fix $w_a \in \mathbb{R}$.

(i) For $\frac{w_a}{w_b} = 1$, $\gamma_1 > 1$ and $\gamma_0 < 0$. Therefore, $(\mu, \mu) \in \overline{D}_h$, since $\mu < 0$.
(ii) For $\frac{w_a}{w_b} \to \infty$, $\gamma_1 \to 1$ and $\gamma_0 > 0$. Therefore $(\mu, \mu) \in \underline{D}_h$.

To complete the argument, it suffices to note that the variation of $\frac{w_a}{w_b}$ from 0 to $+\infty$, corresponds to the combination of two monotone, continuous applications: the rotation of $D_b^{\perp}$ around the axis $\left(\frac{c}{w_a}, \frac{c}{w_b}\right)$ and a translation of $\left(\frac{c}{w_a}, \frac{c}{w_b}\right)$. There exists therefore a unique δ such that for $\frac{w_a}{w_b} = \delta$,

$$(\mu, \mu) \in \underline{D}_h(w_a, w_b) \text{ and } (\mu, \mu) \in \overline{D}_h(w_a, w_b).$$

The rest of the argument follows the same line as the previous proof.

Remark 1. *Assumptions 1.1 and 1.2 can be transformed to accommodate differences in costs and average returns. What is really needed is for the mean* (μ_a, μ_b) *to belong to the lower contour set of the function defined by the plane of symmetry. This will be true, in particular, when the high income-earner has on average better return than the low-income earner.*

Remark 2. *Assumptions 1.3 and 1.4 can be relaxed to allow for a variance matrix of the general form:*

$$\Sigma = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix}$$

The main point is that:

$$\left\| x^{a-} - (\mu_a, \mu_b) \right\|_\Sigma \geq \left\| x^{ab-} - (\mu_a, \mu_b) \right\|_\Sigma \tag{A.6}$$

for all $x^{a-} \in \zeta_{ab}^{a-}$ *and* $x^{ab-} \in \zeta_a^{ab-}$ *and* $\|x\|_\Sigma = x'\Sigma^{-1}x$. *This will be true in particular, when the high income earner has larger variance of return to migration than the low income earner. Again, in general, the class of symmetric distributions will satisfy Eq. (A.6) when Assumptions 1.1 and 1.2 hold.*

Appendix B. Descriptive Statistics

Table 5. Educational distribution for couples and singles

	Females		Males	
	Singles	Couples	Singles	Couples
Low education	29.7%	23.1%	32.0%	21.1%
Lower middle edu.	46.7%	50.4%	51.5%	55.6%
Higher middle edu.	16.2%	19.9%	9.6%	14.7%
High education	7.5%	6.6%	6.9%	8.6%
Observations	2,280,268	4,198,520	3,188,687	4,198,520

Table 6. Probit estimation results for couple and single migration

	Couples	Male Singles	Female Singles
Male 30-34	-0.0258*	-0.0603***	
	(0.0149)	(0.00686)	
Male 35-39	-0.102***	-0.176***	
	(0.0178)	(0.00818)	
Low mid. educ. male	0.141***	0.215***	
	(0.0193)	(0.00815)	
High mid. educ. male	0.395***	0.376***	
	(0.0211)	(0.0107)	
High educ. male	0.637***	0.630***	
	(0.0215)	(0.0101)	
Female 30-34	0.0366***		-0.130***
	(0.0136)		(0.00793)
Female 35-39	-0.00296		-0.357***
	(0.0178)		(0.0103)
Low mid. educ. female	0.114***		0.243***
	(0.0173)		(0.00980)
High mid. educ. male	0.128***		0.268***
	(0.0193)		(0.0120)
High educ. female	0.229***		0.583***
	(0.0218)		(0.0120)
Constant	-3.490***	-2.815***	-2.741***
	(0.0230)	(0.00768)	(0.00914)
Observations	4,198,520	3,188,687	2,280,268
Pseudo R-squared	0.0449	0.0264	0.0300

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

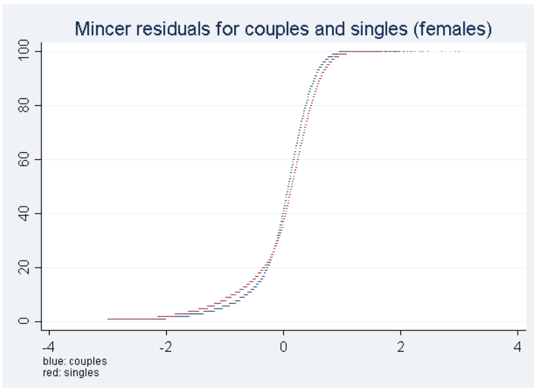

Figure 2. Distribution of earnings residuals, females

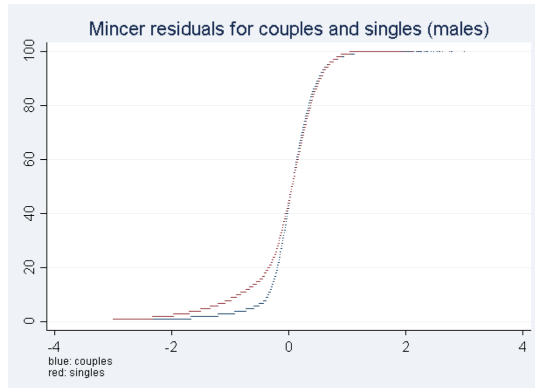

Figure 3. Distribution of earnings residuals, males

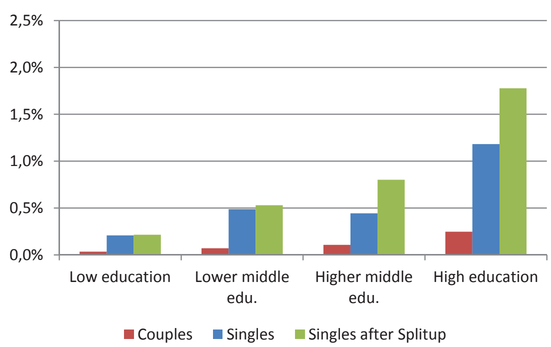

Figure 4. Migration probabilities for females.

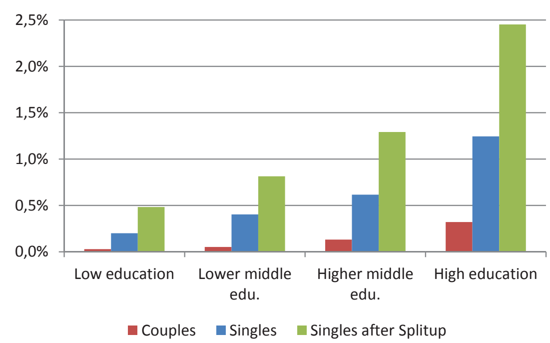

Figure 5. Migration probabilities for males.

Appendix C. Unconditional MIV: A Proposal for an Inference Procedure

We propose a two-stage procedure to retrieve valid confidence regions for the unconditional MIV case, in the event Z is a discrete random variable. The idea is to first apply a precision correction to the term within the min / max operator, before proceeding with the usual CLR procedure with the expectation.

For example, suppose we wish to construct a confidence region of level p for the LB in Eq. (2.4). In the first stage, consider the vector $F^l(z) = \big(F^l(z_1), \dots F^l(z_d)\big)$, where d is the dimension of the support of z. By a first derivation of a precision correction to the vector $F_l(z)$, we obtain a random vector $F^l_{corr}(z)$ with component

$$F^l_{corr}(z_1) \equiv \max_{Z \geq z_1} \left(F(Y=1; M=s|X,Z) - k(p_1)s(X,Z) \right)$$

defined for each z_1 on the support of Z. Recall that $k(p)s(X,Z)$ is a critical value derived from the p-th quantile of an appropriate distribution as described by CLR, times the pointwise standard error. Note that $k(p)$ is the same for each component of the vector.

In the second stage, it suffices now to apply the CLR procedure to the "precision-corrected" variable $E_Z\left(F^l_{corr}(Z)\right)$, to obtain a confidence region of level p_2 for the aforementionned variable.

The first stage ensures that we find a function F^l_{corr} such that asymptotically :

$$\lim_n \mathbb{P}\left(F^l_{corr}(Z) \leq F^l(Z)\right) = 1 - p_1$$

The second stage then ensures that we find a valid confidence region such that asymptotically:

$$\lim_n \mathbb{P}\left(\hat{E}_Z F^l_{corr}(Z) - \tilde{k}(p_2)\tilde{s} \leq E_Z F^l_{corr}(Z)\right) = 1 - p_2$$

Appealing now to the composition theorem from Galichon and Henry (2013) (**Theorem 1**), we find that:

$$\lim_n \mathbb{P}\left(\hat{E}_Z F^l_{corr}(Z) - \tilde{k}(p_2)\tilde{s} \leq E_Z F^l(Z)\right) \geq 1 - p_1 - p_2$$

A similar two-stage procedure is implemented by Méango (2014) to extend the inference procedure proposed by Henry, Méango, and Queyranne (2011) to a case where one of the dependent variables is censored. A thorough investigation of the finite sample properties of the proposed procedure is beyond the scope of this paper, however.

Appendix D. Unconditional MIV: Simulations for a Two-Stage CLR Procedure

D.1. **Case Study 1.** We generate the following variable Y, D, U such that:

$$\begin{aligned} U &\sim Unif([0,1])^2 && \text{(D.1)} \\ D &= I(0.1 + Z > U_1) && \text{(D.2)} \\ Y &= I(0.6 + 0.5D + \beta Z > U_2) && \text{(D.3)} \end{aligned}$$

where Z is a multinomial random variable that has support $\{0, 1, 0.2, ..., 0.5\}$. Each element of the support occurs with equal probability. We approximate the distribution of the true data generating process for the following value of $\beta = 0, 0.1, 0.5$, for 2×10^8 draws of the random variables U and Z. This distribution is summarized in the three first columns of Table 7. Note that $P(Y = 1|Z)$ is strictly increasing in Z in accordance with the MIV assumption. Most importantly, we have:

$$E_Z \max_{Z \leq z_0} P(Y = 1; D = 0|Z = z_0) \ \leq \ P(Y = 1|D = 0)$$

Note also that $P(Y = 1; D = 0|Z = z_0)$ is strictly decreasing in z_0, so that the unconditional MIV lower bound is given by

$$E_Z \max_{Z \leq z_0} P(Y = 1; D = 0|Z = z_0) = \max_{Z} P(Y = 1; D = 0|Z = z) = P(Y = 1; D = 0|Z = 0.1). \tag{D.4}$$

We compute a one-sided confidence region of level p for the unconditional MIV bound given by

$$E_Z \max_{Z \leq z} P(Y = 1; D = 0|Z = z)$$

using two methods: (i) the two-stage CLR procedure advised in the main text, say TSCLR, and for comparison (ii) a "naïve" implementation of the CLR bounds. The latter procedure is conducted as follow: first, compute $\hat{P}(Y = 1; D = 0|Z = z)$. Then take the max of the estimator on the appropriate set for each z, on the support of z. For example, for $z_0 = 0.3$, the set is given by $\{0.1, 0.2, 0.3\}$. Compute the estimator $\hat{\theta}_{naive} = \sum_z \max \hat{P}(Y = 1; D = 0|Z = z) \times \hat{P}(Z)$. The one-sided confidence region is given by the term: $\hat{\theta}_{naive} - \tilde{k}(p)\hat{s}_{naive}$, where $\hat{s}_{naive}$ is the standard error of $\hat{\theta}_{naive}$ and $\tilde{k}(p)$ is an estimated critical value, as described in Chernozhukov, Lee, and

Table 7. Distribution of the true data generating process

	Case Study 1			Case Study 2		
Values for the parameter β	0	0.1	0.5	0	0.1	0.5
$P(Y=1; D=0)$	0.36	0.376	0.44	0.5259	0.53	0.5313
$P(Y=1\|D=0)$	0.6	0.6266	0.7332	0.8414	0.8479	0.8499
$P(Y=1; D=0\|Z=z_0)$	0.4799	0.4879	0.5198	0.0116	0.0109	0.0069
$P(Y=1; D=0\|Z=z_1)$	0.4199	0.434	0.4898	0.6379	0.6186	0.5242
$P(Y=1; D=0\|Z=z_2)$	0.3601	0.3779	0.45	0.8223	0.8223	0.8223
$P(Y=1; D=0\|Z=z_3)$	0.3001	0.3201	0.4001	0.8037	0.8257	0.8914
$P(Y=1; D=0\|Z=z_4)$	0.2402	0.2601	0.34	0.3541	0.3725	0.4113
$E_Z \max_{Z\leq z} P(Y=1; D=0\|z)$	0.4799	0.4879	0.5198	0.6233	0.6206	0.6272

Rosen (2013). For each sample size, $n = 100$, $n = 500$, and $n = 5,000$ we simulate

Table 8. Case study 1: Monte-Carlo simulations results

		TSCLR$< E_Z F^l$			naïve CLR $< E_Z F^l$			TSCLR $> CLR(P_10)$		
β	Level (α)	$n = 100$	500	5000	100	500	5000	100	500	5000
0	0.95	1	1	1	0.982	0.987	0.997	0.02	0.563	1
	0.9	0.998	1	1	0.962	0.975	0.978	0.038	0.655	1
	0.5	0.97	0.975	0.971	0.734	0.729	0.748	0.212	0.932	1
0.1	0.95	1	1	1	0.976	0.988	1.985	0.019	0.48	1
	0.9	0.998	0.999	1	0.953	0.97	0.962	0.028	0.567	1
	0.5	0.967	0.97	0.962	0.719	0.711	0.686	0.181	0.899	1
0.5	0.95	0.999	0.999	1	0.963	0.967	0.921	0.012	0.212	1
	0.9	0.998	0.998	0.998	0.93	0.932	0.848	0.02	0.289	1
	0.5	0.957	0.949	0.856	0.687	0.637	0.412	0.131	0.675	1

1,000 samples[18]. We use $1,000$ draws from a normal random distribution within the implementation of the CLR procedure to compute the critical value. We consider confidence levels 95%, 90%, and 50%. Note that for the TSCLR, we define the level

[18]The intermediate size $n = 1000$ gives results to similar to those found when $n = 500$. We chose therefore not to report the results

of the first step as $p_1 = 97.5\%, 95\%, 50.56\%$ and the second step accordingly. Coverage probabilities of the identified set by the confidence region, as computed from the 1,000 samples for the TSCLR procedure and the naïve procedure are displayed in the first and second group of columns of Table 8, respectively. In the last group of columns, we show the frequency with which the bound obtained from the $TSCLR$ procedure is larger than the one obtained by ignoring the unconditional MIV assumption.

The TSCLR procedure is conservative. Monte Carlo frequency of rejection by using the TSCLR procedure is lower in all cases than the theoretical level. This over-rejection is corrected for by an increase in the sample size only in the case where $\beta = 0.5$. Fig.6 shows the distribution of the TSCLR critical value for different sample sizes. With small sample sizes, the distribution is more spread out and shifted to the left, toward lower values. A larger sample size progressively creates a more concentrated distribution, with a higher median.

The naïve procedure exhibits less conservative coverage properties. Indeed, the distribution of the TSCLR critical value is first-order stochastically dominated by the critical value obtained from the naïve procedure (see Fig. 7). This is the result of the first-step correction. Note however, that Monte Carlo frequency of rejection in the naïve procedure is higher than the theoretical level for $\beta = 0.5$ and the largest sample size. Note also that the large discrepancy in coverage between the two procedures mainly originates from the fact that the unconditional MIV minorant is almost surely a constant given Z by Eq. (D.4). The second step correction of the TSCLR procedure could then be seen as an over-correction.

Finally, with small sample size, $n = 100$, the critical values obtained by ignoring the unconditional MIV assumption produce tighter confidence regions than those obtained through the TSCLR procedure in at least 78% of the Monte-Carlo replications ($\beta = 0, \alpha = 0.5$) and as much as 99% of the replications ($\beta = 0.5, \alpha = 0.99$). This measure of the intrinsic usefulness of the TSCLR improves significantly with sample size. With the largest sample size, the confidence region obtained is always tighter when using the TSCLR procedure.

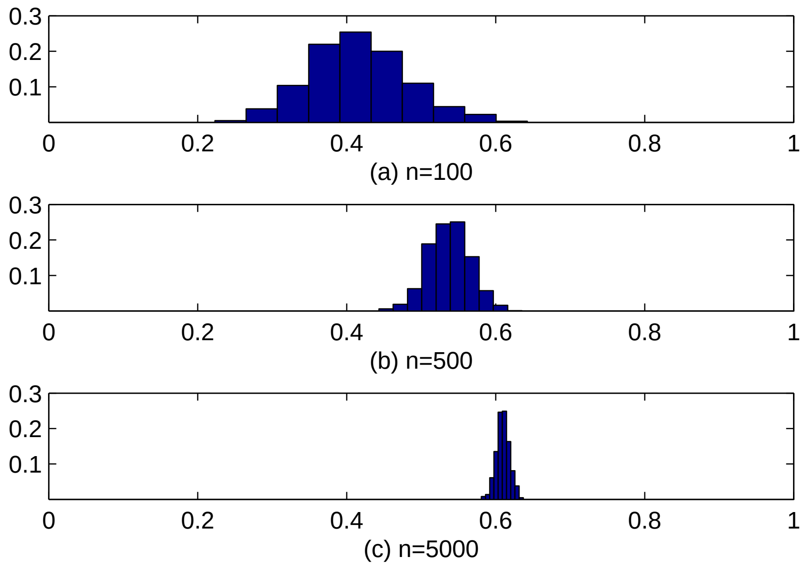

Figure 6. Distribution of the critical value obtained by applying the TSCLR procedure.

The distribution is obtained for Case Study 2, the level being fixed at 0.90 and $\beta = 0$. Panel (a) ((b), (c)) shows the Monte Carlo frequency for 1,000 samples of size $n = 100$ (500, 5,000). The distribution is more spread out and shifted to the left with small sample size (n=100). A larger sample size progressively creates a more concentrated distribtution, with higher median.

D.2. **Case study 2.** We generate the following variable Y, D, U such that:

$$V \sim N([0,1])^2 \tag{D.5}$$

$$D = I(-2 - 0.5Z + 0.8Z^2 > V_1) \tag{D.6}$$

$$Y = I(1 + 0.1D + \beta Z > V_2) \tag{D.7}$$

where Z is a multinomial random variable that has support $\{-0.2, -0.1, \ldots, 0.2\}$. Each element of the support occurs with equal probability. We approximate the distribution of the true data generating process for the following value of $\beta = 0, 0.1, 0.5$, for 2×10^8 draws of the random variables V and Z. This distribution is summarized in the last group of columns of Table 7. Note that $P(Y = 1; D = 0|Z = z)$ is now

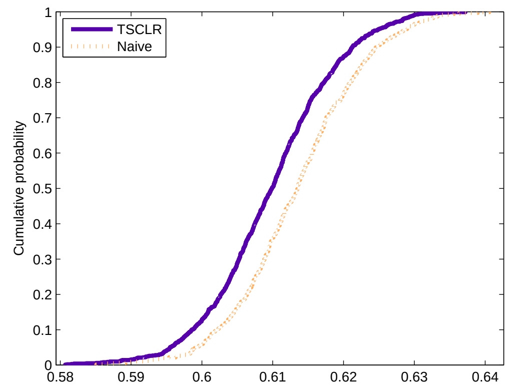

FIGURE 7. Cumulative Distribution of the critical value obtained by applying the TSCLR procedure and the naïve procedure.
The cumulative distribution is obtained for the Case study 2, the level being fixed to 0.90, $\beta = 0$ and $n = 5000$. The distribution of the TSCLR critical value is first-order stochastically dominated by the one of the critical value obtained from the naïve procedure

first increasing in z, then decreasing in Z, so that

$$E_Z \max_{Z \leq z} P(Y = 1; D = 0 | Z = z) \neq \max_{Z} P(Y = 1; D = 0 | Z = z). \tag{D.8}$$

TABLE 9. Case study 2: Monte-Carlo simulations results

		TSCLR$< E_Z F^l$			naïve CLR$< E_Z F^l$			TSCLR $> CLR(P_1 0)$		
β	Level (α)	$n = 100$	500	5000	100	500	5000	100	500	5000
0	0.95	0.998	1	0.97	0.996	0.999	0.943	0.093	0.981	1
	0.9	0.996	0.998	0.932	0.996	0.997	0.867	0.131	0.992	1
	0.5	0.97	0.949	0.546	0.904	0.922	0.215	0.246	0.999	1
0.1	0.95	0.999	1	0.969	0.999	0.999	0.942	0.074	0.943	1
	0.9	0.998	0.998	0.938	0.995	0.997	0.85	0.102	0.966	1
	0.5	0.976	0.949	0.556	0.883	0.779	0.204	0.209	0.992	1
0.5	0.95	1	1	1	0.996	1	1	0.036	0.69	1
	0.9	1	1	1	0.996	0.999	1	0.052	0.787	1
	0.5	0.992	0.993	1	0.904	0.922	0.978	0.103	0.904	1

The TSCLR procedure is found to be conservative. Compared to the previous case though, improvement of the coverage rate occurs with an increase of sample size in the case where $\beta = 0, 0.1$. In this case, the Monte Carlo frequency of rejection is relatively close to the theoretical level. Note that in comparison, the naïve procedure exhibits higher Monte Carlo frequency of rejection than the theoretical level. Finally, the TSCLR is more often informative than the CLR procedure ignoring the unconditional MIV assumption, even for a sample size of $n = 500$.

Overall, the procedure is conservative and this aspect might (or might not) be corrected with the increasing sample size. Even so, however, the procedure may retain some information contained in the unconditional MIV assumption and tighten the confidence region. Ideally, the procedure should be coupled with a test of the Condition D.4, so as to avoid a overcorrection. Note also that the AIS correction from Chernozhukov, Lee, and Rosen (2013) was not applied in the simulation and might improve the coverage properties of the procedure. We leave this to future research.