

SHARP BOUNDS AND TESTABILITY OF A ROY MODEL OF STEM MAJOR CHOICES

ISMAËL MOURIFIÉ, MARC HENRY, AND ROMUALD MÉANGO

ABSTRACT. We analyze the empirical content of the Roy model, stripped down to its essential features, namely sector specific unobserved heterogeneity and self-selection on the basis of potential outcomes. We characterize sharp bounds on the joint distribution of potential outcomes and testable implications of the Roy self-selection model under an instrumental constraint on the joint distribution of potential outcomes we call *stochastically monotone instrumental variable (SMIV)*. We show that testing the Roy model selection is equivalent to testing stochastic monotonicity of observed outcomes relative to the instrument. We apply our sharp bounds to the derivation of a measure of departure from Roy self-selection to identify values of observable characteristics that induce the most costly misallocation of talent and sector and are therefore prime targets for intervention. Special emphasis is put on the case of binary outcomes, which has received little attention in the literature to date. For richer sets of outcomes, we emphasize the distinction between pointwise sharp bounds and functional sharp bounds, and its importance, when constructing sharp bounds on functional features, such as inequality measures. We analyze a Roy model of college major choice in Canada and Germany within this framework, and we take a new look at the under-representation of women in STEM.

Date: The first version is of 22 April 2012. The present version is of April 2, 2019. This research was supported by SSHRC Grants 410-2010-242, 435-2013-0292 and 435-2018-1273, NSERC Grant 356491-2013, and Leibniz Association Grant SAW-2012-ifo-3. The research was conducted in part, while Marc Henry was visiting the University of Tokyo and Ismaël Mourifié was visiting Penn State and the University of Chicago. The authors thank their respective hosts for their hospitality and support. They also thank Désiré Kédagni, Lixiong Li, Karim N'Chare, Idrissa Ouili and particularly Thomas Russell and Sara Hossain for excellent research assistance. Helpful discussions with Laurent Davezies, James Heckman, Hidehiko Ichimura, Koen Jochmans, Essie Maasoumi, Chuck Manski, Ulrich Müller, Aureo de Paula, Azeem Shaikh and helpful comments from five anonymous referees, and from seminar audiences in Cambridge, Chicago, École polytechnique, Emory, Indiana, Northwestern, Penn State, Princeton, Toronto, Toulouse, UCL, UPenn, Vanderbilt and the participants of the CESG and the 2018 Canadian senate open caucus on women and girls in STEM are also gratefully acknowledged. Correspondence address: Department of Economics, Max Gluskin House, University of Toronto, 150 St. George St., Toronto, Ontario M5S 3G7, Canada.

Keywords: Roy model, sectorial choice, partial identification, stochastic monotonicity, intersection bounds, functional sharp bounds, inequality, optimal transport, returns to education, college major, gender profiling, STEM, SMIV.

JEL subject classification: C31, C34, C35, I21, J24

INTRODUCTION

In a seminal contribution that is now part of the folklore of economics, Roy (1951) proposed a model of earnings with sorting on sector specific skills. Roy's objective was to provide a channel by which skills translate into earnings and to capture the idea that favorable sorting reduces earnings inequality. The simplicity of this mechanism and the richness of its implications turned the Roy model into one of the most successful tools in the analysis of environments, where skills and choices interact: they include the Gronau-Heckman (1974) labor supply model, the unionization model of Lee (1978), the model of education self-selection proposed by Willis and Rosen (1979), sector selection in Heckman and Sedlacek (1985), and the Borjas (1987) immigration model. More recently, Lemieux (1998) revisited the issue of inequality in the unionization model, Mulligan and Rubinstein (2004) used the Roy model to shed light on the recent evolution of the gender gap, Chandra and Staiger (2007) to analyze the choice of surgical procedures and Eisenhauer, Heckman, and Vytlačil (2015) to analyze benefits and costs of educational choices. The list is, of course, far from complete, but quite sufficient to show the enormous success of the Roy model.

In the original model, skills are jointly log normal and Basu and Ghosh (1978) show that under this assumption, the joint distribution of skills and the marginal distributions of potential earnings are identified. Heckman and Honoré (1990) further show that self-selection does indeed reduce aggregate inequality when skills are log normal and within sector inequality when skills have a log concave distribution. Naturally, the effect of self-selection on outcome inequality remains empirically relevant when skills do not have a log concave distribution. However, the analysis of the nonparametric version of the Roy model, stripped down to the self-selection mechanism, has long been hampered by (lack of) identification issues. The Cox (1972) and Tsiatis (1975) comments on non-identifiability of competing risks imply that any continuous outcome distribution could be rationalized with independent sector-specific skills, so that the Willis and Rosen (1979) notion of skill hierarchies loses empirical content.

One way to resolve this lack of identification issue, pioneered by Heckman and Honoré (1989, 1990), is to bring in additional information to achieve identification, such as repeated cross sections, panel data, factor structure, exclusion restrictions and large support assumptions within restricted specifications of the model. A vast literature, both theoretical and empirical, followed this lead (see for instance Heckman (2001), Heckman and Vytlačil (2007), Heckman and Taber (2008) or French and Taber (2011) for recent accounts). Another way to approach the issue, which was pioneered by Peterson (1976) and which we follow here, is to recognize that, despite the identification failure because of self-selection, the Roy model is not devoid of empirical content.

The object of the present article is to characterize this empirical content, with special emphasis on the joint distribution of potential outcomes and testability of the Roy selection mechanism. This implies considering distributional features of outcomes, which are important if one is to evaluate

the effect of self-selection on wage inequality, as Roy initially intended; see Cunha, Heckman, and Navarro (2006) for a discussion. It further implies considering joint distributional aspects. As Heckman (1992) noted, information on the joint distribution is necessary to evaluate welfare implications of policy changes that affect the relative price of skills in both sectors. Correlation between outcomes can be important to policy evaluation, as discussed in Carneiro, Hansen, and Heckman (2002), as can the difference between potential outcomes or the distribution of outcomes conditional on the chosen sector. In all such cases, the joint distribution of potential outcomes is the relevant object to characterize. We refer to Heckman (2010), Heckman, Smith, and Clements (1997), and Abbring and Heckman (2007) for in-depth discussions of this issue.

We devote a considerable amount of attention to the case with binary outcomes, which we call the *binary outcomes Roy model*. The reason is twofold. First, the identification failure is starker with binary outcomes, and the characterization of the joint distribution is easier to derive and explain in the binary case, before it is extended to the more general cases of discrete, continuous or mixed discrete and continuous outcomes. Second, the case of discrete outcomes has received very little attention in the Roy model literature. Most of this literature concerns the case of continuous outcomes and many applications, where outcomes are discrete, fall outside its scope. They include analysis of the effects of different training programs on the ability to secure employment, of competing medical treatments or surgical procedures on survival, of higher education on migration and of competing policies on schooling decisions in developing countries among numerous others. The Roy model is still highly relevant to those applications, but very little is known of its empirical content in such cases.

We derive sharp bounds for the joint distribution of (Y_0, Y_1) , using techniques from Galichon and Henry (2011) and Beresteanu, Molchanov, and Molinari (2011) -see also Beresteanu, Molchanov, and Molinari (2012) and Chesher, Rosen, and Smolinski (2013). Bounds do not cross, and the model is not testable, unless we observe variables that have a restricted effect on potential outcomes. A special case is that of selection shifters that are statistically independent of potential outcomes. Such variables have two major drawbacks in this framework. First, they are very elusive in important areas of application of this methodology. To take one classical example in the literature on returns to education, parental education, measures of school quality, and fees may be correlated with unobserved cognitive and non cognitive parental investments and is therefore unlikely to be independent of potential outcomes. Second, within the Roy model, given the sector selection mechanism, a variable that is independent of potential outcomes can only affect sector selection if potential outcomes are equal in both sectors, which severely restricts the extent of resulting variation in sector selection.

To resolve both of these issues, we introduce *stochastically monotone instrumental variables*. They are selection shifters that are restricted to affect potential outcomes monotonically. For instance, parental education may not be independent of potential wages because of unobserved parental cognitive and non cognitive investments. However, it is unlikely that such additional investments

will negatively affect potential future wages. Moreover, allowing for monotonic effects on potential outcomes resolves the second issue, since stochastically monotone instrumental variables may shift selection even when potential outcomes are different. Our stochastically monotone instrumental variable assumption is stronger than the Manski and Pepper (2000) *monotone instrumental variable* assumption, which only requires mean potential outcomes to be monotonic in the instrument, rather than the whole distribution, as we do. This difference is crucial, when deriving bounds on (joint) distributional features, rather than simply averages.

We derive the identified set for the joint distribution of potential outcomes in the binary outcomes Roy model under this assumption of stochastically monotone instrumental variable (hereafter *SMIV*), and we show that stochastic monotonicity of observed outcomes in the instrument summarizes all observable implications of the model. Hence a test of Roy selection behavior boils down to a test of stochastic monotonicity, and can therefore be conducted with existing inference methods, as in Lee, Linton, and Whang (2009), Delgado and Escanciano (2012) and Hsu, Liu, and Shi (2016). This provides a fully nonparametric alternative to tests of the Roy model proposed since Heckman and Sedlacek (1990) (see also Dahl (2002), with multiple treatments and repeated cross-sections). Intuition about the relation between stochastic monotonicity and Roy selection can be gained from the following hypothetical scenario. Suppose two identical young women have higher economic prospects in non STEM fields. One of them is induced to choose a STEM degree by a larger proportion of women on the STEM faculty in her region, whereas the other, who lives in a region with lower proportion of women on the STEM faculty, chooses a non-STEM field to maximize economic prospects. It will then appear that an increase in the proportion of women on the STEM faculty produces a decrease in observed outcomes, hence a rejection of monotonicity.

To alleviate the concern that rejection of Roy selection behavior may in fact be down to a rejection of the assumption that individuals are perfectly informed of their potential future outcomes at the time of sector selection, we derive sharp bounds for the joint distribution of potential outcomes for an imperfect foresight binary outcomes Roy model, where agents select the sector that maximizes the expectation of their outcome with respect to their information set at the time of decision. Since the model is rejected if and only if this identified set is empty, it allows us to summarize all observable implications of that version of the model as well. We also provide bounds for a measure of departure from Roy selection, which is constructed from the difference between the maximum potential outcome and the realized outcome (both being equal under Roy selection), and which, again, requires bounding the joint distribution of potential outcomes, as we do here, rather than marginal average outcomes, as is customary in the literature. These measures of departure from the Roy selection model serve to identify values of observable characteristics that induce the most costly misallocation of talent and field of study and are therefore prime targets for intervention.

We extend the test of Roy self-sorting and the measure of departure from Roy to more general discrete and continuous outcomes. When extending the analysis of the bounds on potential outcomes,

distributional issues come to the fore. The classic Peterson (1976) bounds are sharp for $\mathbb{P}(Y_0 \leq y)$ and $\mathbb{P}(Y_1 \leq y)$, for each quantile y , but, as noted by Crowder (1991), they do not incorporate monotonicity and right-continuity restrictions on distribution functions. Hence they entail loss of information, when the object of interest involves densities, such as hazard rates, or functionals of the distribution, such as inequality measures. We provide a general characterization of the joint distribution and bounds on the marginal distributions of potential outcomes that are *functionally sharp*, in the sense that they incorporate slope restrictions. In this, we follow Bedford and Meilijson (1997), although the model specification, hence the bounds, are different. We apply the latter bounds to derive sharp bounds on inequality measures, which we show are more informative than would have been obtained from *pointwise sharp* bounds on the distributions of potential outcomes, such as Peterson bounds.

In the tradition of Willis and Rosen (1979), Kenny, Lee, Maddala, and Trost (1979) and, more recently, Eisenhauer, Heckman, and Vytlačil (2015), we analyze returns to education through the lens of the Roy selection model. It is well documented, since at least James, Alsalam, Conaty, and To (1989), that major choice is an important determinant of labour market outcomes, and the account in Altonji, Arcidiacono, and Maurel (2016) shows that the literature on the determinants of and the returns to major choice is now substantial. The STEM versus non-STEM classification has come to dominate the debate. We therefore analyze a Roy model of choice of field of study, between STEM and non-STEM degrees, based on data from both the Canadian and German National Graduate Surveys (NGS). Following the recent literature on the subject, surveyed in Kahn and Ginther (2017), we focus on mathematics intensive fields, including economics, but excluding life sciences.

We consider a Roy model of major choice, where the target labour market outcomes are wage, obtaining a permanent job within a year of graduation, or holding a job related to the field of study, respectively. Our main objective is to shed some insight onto the under-representation of women in STEM education and even more so in STEM jobs, the gender gap in STEM labour market outcomes, and the contribution of the STEM economy to rising wage inequality. If choices conform to the Roy self-sorting mechanism, only policies directed at ex-post wage discrimination are likely to be effective in reducing inefficiencies, not policies directed at reducing gender profiling in major choice. Hence, in our investigation of the determinants of under-representation of women in STEM fields, we give prominence to testing Roy selection behavior. As the latter is not testable without covariate restrictions, we consider variation induced in major choices by parental education. Although there are reasons to doubt the validity of parental education according to Assumption 4 below, which requires independence of the instrument and the vector of potential outcomes, it is reasonable to assume parental education level has a monotonic effect on potential outcomes, as prescribed by Assumption 5 below, which requires stochastic monotonicity of the distribution of potential outcomes, conditionally on the instrument. When testing whether women graduate choices conform to the Roy self-sorting mechanism, we also use the fraction of women in the faculty

of STEM programs in the region and at the time of major choice as an instrument, based on the assumption that role models may not negatively affect future prospects for women graduates. We also perform our test of Roy self-sorting based on additional vectors of instruments incorporating local labor market conditions at the time of choice for robustness purposes, and find very little variation in test results.

Our tests of the Roy model for different choices of instruments and different employment related outcome variables reveal significant gender, racial and regional differences. We find a pattern of rejections of Roy self-sorting based on outcomes for white women in the former Federal Republic of Germany and the rest of Canada, and a lack of rejections for visible minorities and for white women from Québec, white men from all of Canada and the former German Democratic Republic. Confidence intervals for measures of departure from Roy behavior reveal that in the case of white women from the former Federal Republic, for instance, rejection of Roy behavior seems to be driven by lower income women with high school educated mothers and middle income women with postgraduate educated mothers. Among groups, where Roy self-sorting is not rejected, comparisons of interquartile ranges for observed and counterfactual income distributions are inconclusive except in the case of white women in the former Democratic Republic and white men in Québec, where self-sorting is found not to increase inequality, complementing the result for log-concave talent distributions in Heckman and Honoré (1990).

The pattern of rejections of Roy self-sorting in major choice points to non labor market related determinants of choice. For instance, our results are consistent with a story involving gender profiling pushing white men in the West of Germany into STEM fields and white women in the West of Germany and in Canada out of STEM fields. They are also consistent with gender profiling being less prevalent in the former communist Germany. However, the results are also consistent with a story involving non pecuniary field preferences driving major choices of more privileged groups in more affluent regions, but not the choices of the more financially constrained.

Other related literature.

Sharp bounds are derived in binary outcome models with a binary endogenous regressor in Chesher (2010), Shaikh and Vytlacil (2005, 2011), Chiburis (2010), Jun, Pinkse, and Xu (2010) and Mourifié (2011) under a variety of assumptions, which all rule out sector specific unobserved heterogeneity. Aakvik, Heckman, and Vytlacil (2005) provide identification conditions for a version of the binary outcomes Roy model, where sector specific unobserved heterogeneity follows a factor structure. Chesher, Rosen, and Smolinski (2013) derive sharp bounds for instrumental variable models of discrete choice. Balke and Pearl (1997) propose sharp bounds for the average treatment effect with binary endogenous treatment and a binary instrument. Manski and Pepper (2000) and Heckman and Vytlacil (2001) refine them with additional monotonicity assumptions and extend them beyond binary instruments.

Recent developments in nonparametric inference in Roy and competing risks models can be found in Khan and Tamer (2007), Lee and Lewbel (2013), Bayer, Khan, and Timmins (2011), D’Haultfoeuille and Maurel (2013) and references therein. Heckman, Smith, and Clements (1997), and later Firpo and Ridder (2008), Fan and Park (2010) and Fan and Wu (2010) look at more general treatment effects models from a partial identification point of view and use rearrangement inequalities to derive bounds on the distribution of treatments effects, a feature of the joint distribution of potential outcomes, under conditions, where the marginal distributions of potential outcomes are identified.

Caplin and Nalebuff (1991) derive bounds on earnings inequality based on the Mandelbrot (1962) factor model version of the Roy mechanism. Sharp bounds on inequality measures for the potential outcome distributions are discussed in Vasquez-Alvarez, Melenberg, and van Soest (2003), Blundell, Gosling, Ichimura, and Meghir (2007) and Stoye (2010). Blundell, Gosling, Ichimura, and Meghir (2007) look specifically at the interquantile range, while Stoye looks at a more general family of measures. Blundell, Gosling, Ichimura, and Meghir (2007) and Stoye (2010) consider general models with treatment selection on unobservables. In the case of the Roy model, we encounter the same issue that pointwise bounds on the potential distributions do not deliver sharp bounds on the interquantile range.

Arcidiacono (2004) and Befy, Fougère, and Maurel (2012) are among the recent investigations of major choice. Kirkebøen, Leuven, and Mogstad (2016) use information on the ordering of non chosen options in a multi-sector treatment model under traditional IV assumptions. They find evidence of behavior consistent with pure Roy sorting on income for most pairs of preferred and next preferred field choices. More references can be found in the survey of research on heterogeneous human capital investments by Altonji, Blom, and Meghir (2012). The gender gap in STEM labour market outcomes is highlighted by Zafar (2013), Hunt, Garant, Herman, and Munroe (2013) and Card and Payne (2017). We follow Boudarbat and Montmarquette (2007) and Lemieux (2015) in the use of Canada’s National Graduate Survey to study determinants and consequences of the choice of major.

Outline. The remainder of the paper is organized as follows. Section 1 details the general frame of analysis. Section 2 concerns sharp bounds and testability of the binary outcomes Roy model, and its version with imperfect foresight, under the assumption of stochastic monotonicity of potential outcomes relative to an instrument. Section 3 derives functional sharp bounds for the Roy model with mixed discrete-continuous outcomes, and their implications for testability of Roy behavior and effects of endogenous sector selection on functional features such as inequality measures. Section 4 applies the derived bounds to the analysis of major choice in Canada and Germany and the under-representation of women in STEM. The last section concludes. Proofs of the main results are collected in the appendix.

1. ANALYTICAL FRAMEWORK

We adopt the framework of the potential outcomes model $Y = Y_1D + Y_0(1 - D)$, where Y is an observed scalar outcome, D is an observed selection indicator, which takes value 1 if Sector 1 is chosen, and 0 if Sector 0 is chosen, and Y_1, Y_0 , are unobserved potential outcomes, with common lower bound \underline{b} on their support (\underline{b} is usually 0 or $-\infty$). Heckman and Vytlacil (1999) trace the genealogy of this model and we refer to them for terminology and attribution. The object of interest is the joint distribution of (Y_0, Y_1) and features thereof. Since Y and D are observed, the joint distribution of (Y, D) is directly identified from the data.

We strip the model down to its self-selection mechanism, where agents are perfectly informed of the joint distribution of their potential outcomes (Y_0, Y_1) in both sectors and choose the sector that maximizes outcomes, so that $D = 1$ when $Y_1 > Y_0$ and $D = 0$ if $Y_1 < Y_0$. The model is silent on the tie-breaking mechanism agents use in case $Y_1 = Y_0$. As is customary in such frameworks, the assumption that agents are perfectly informed is intended to reflect, within a simple static model, the result of dynamic adjustments and learning on the one hand, and to put in stark relief the difference between the agents' and the analyst's information sets, on the other hand.

We summarize the model with the following assumptions.

Assumption 1 (Potential outcomes). *Observed outcomes are the realizations of a random variable Y satisfying $Y = Y_1D + Y_0(1 - D)$, where (Y_0, Y_1) is a pair of possibly dependent unobserved random variables and D is an observed indicator variable.*

Assumption 2 (Selection). *The selection indicator satisfies $Y_1 > Y_0 \Rightarrow D = 1$, $Y_1 < Y_0 \Rightarrow D = 0$.*

Individuals choose the sector that yields higher outcome, when $Y_1 \neq Y_0$. Their choice criterion is unspecified if $Y_1 = Y_0$. When outcomes are discrete, the possibility of ties has to be considered. More generally, in a Roy model of earnings, the possibility of equal earnings in both sectors has to be entertained, if wage setters propose contracts that pool different skill levels, for instance. If the probability of ties is non zero, the Roy model specification described here is different from the specification of traditional competing risks models, contrary to the generally held belief that Roy and competing risks models are identical, up to sign convention. The model we consider here is tuned to economic applications, where the sector selection is unknown, when both sectors yield the same outcome. Hence, we identify $\mathbb{P}(Y_d \leq y, D = d)$, for $d = 0, 1$, but not $\mathbb{P}(Y_d \leq y, Y_d > Y_{1-d})$. All we know is that $\mathbb{P}(Y_d \leq y, Y_d > Y_{1-d}) \leq \mathbb{P}(Y_d \leq y, D = d) \leq \mathbb{P}(Y_d \leq y, Y_d \geq Y_{1-d})$. In the competing risks analysis of Bedford and Meilijson (1997), on the other hand, $\mathbb{P}(Y_1 \leq y, Y_1 > Y_0)$, $\mathbb{P}(Y_0 \leq y, Y_1 < Y_0)$ and $\mathbb{P}(Y_1 \leq y, Y_1 = Y_0)$ are all assumed identified, so that one observes when both components of the system fail simultaneously.

Remark 1. *Our analysis can be extended to the case, where Y, Y_0 and Y_1 , take values in an ordered subset of a Euclidean space, such as \mathbb{R}^2 endowed with the lexicographic order \lesssim_{lex} , for instance. In*

the latter case, with outcome variable $Y = (W, T)$ ordered lexicographically, Assumption 2, would read $[W_d > W_{1-d} \text{ or } [W_d = W_{1-d} \text{ and } T_d > T_{1-d}]] \Rightarrow Y = Y_d$.

Take the case of university STEM major choice for instance. Lexicographic Roy preferences based on relatedness and income implies that prospective students choose STEM if they anticipate only STEM degrees will provide them with employment in their field of study, or if both STEM and non STEM provide them with employment in their field but they anticipate higher earnings in STEM.

All results below would relate to the probability distributions of outcomes and potential outcomes relative to the chosen order, not the multivariate probability distributions. For instance, in the lexicographic example, the probability distribution is defined as $\mathbb{P}(Y \succ_{lex} y) = \mathbb{P}((W, T) \succ_{lex} (w, t)) = \mathbb{P}(W < w \text{ or } [W = w \text{ and } T \leq t])$.

Remark 2. The whole analysis is understood to be conditional on a set of observed covariates, which will be omitted from the notation, unless they are involved in identifying assumptions.

2. BINARY OUTCOME ROY MODEL

A great deal of the intuition for the characterization that we propose for the Roy model can be developed with the simplest version, where Y_0 and Y_1 are both binary outcomes. It models success or failure in securing a desired outcome, and the way it depends on a binary choice of treatment. In the case of college major choice, considered in Section 4, Y_1 will model the ability to secure permanent employment at the time of the survey interview, if the degree or the major is classified as STEM, whereas Y_0 will model the ability to secure employment, with a non-STEM degree or major.

Definition 1 (Binary outcome Roy model). *A model satisfying Assumptions 1 and 2, with $Y_0, Y_1 \in \{0, 1\}$, is called binary outcome Roy model.*

An alternative way of defining a binary outcomes model, which shares the main features of the Roy model, i.e., self-selection on unobserved heterogeneity, involves *latent potential outcomes*. It is identical to the Roy model, except that potential outcomes are censored.

Definition 2 (Alternative binary Roy model). *Observed outcomes are the realizations of a random variable Y satisfying $Y = Y_1 D + Y_0(1 - D)$, where*

- (1) *potential outcomes satisfy $Y_d = 1\{Y_d^* > 0\}$, for $d = 0, 1$, for a pair of possibly dependent unobserved random variables (Y_0^*, Y_1^*) ,*
- (2) *D is an observed indicator variable, satisfying $Y_1^* > Y_0^* \Rightarrow D = 1$, $Y_1^* < Y_0^* \Rightarrow D = 0$.*

The alternative binary Roy model of Definition 2 can be interpreted in two ways. First, it is equivalent to a model with $Y = 1\{Y^* > 0\}$, where Y^* satisfies a Roy model. Hence, it can be interpreted as a censored Roy model. The latent variables may be continuous variables, such as

wages, and the analyst only observes whether or not they fall above or below a threshold. Other examples include examination grades, which are unobserved, except for the pass or fail outcome. Second, the actual outcome may be binary and be the result of a two-stage decision by the agent. In a first stage, they choose the sector of activity, with their choice of college major, for instance. In a second stage, they decide whether or not to work. The labor supply decision hinges on the difference between wage and reservation wage in the chosen sector. Then, Y_d^* can be interpreted as the difference between wage in Sector d and reservation wage in Sector d . If reservation wages are equal in both sectors, the model still conforms to the simple Roy incentive mechanism, where wages are the only determinant of sector choice. If reservation wages differ in both sectors, however, the model no longer conforms to the simple Roy incentive mechanism, as sector selection internalizes possibly non-pecuniary costs and benefits of each sector, as in the recent analysis of the generalized Roy model in Eisenhauer, Heckman, and Vytlačil (2015).

Despite their distinct interpretations, it will be shown that sharp bounds for the joint distribution of (Y_0, Y_1) are identical in both models, so that both models carry exactly the same information on the joint distribution of censored potential outcomes. They also share the reduced form implication

$$\mathbb{E}(Y_d - Y_{1-d} | D = d) \geq 0, \quad (2.1)$$

which can be interpreted as a condition of *chosen sector advantage* or as nonnegative average treatment on the treated (where choice of Sector 1 corresponds to treatment). However, we show below that the reduced form condition (2.1) contains less information on the joint distribution of potential outcomes than the structural models of Definitions 1 and 2 do. In particular, constraint (2.1) is also shared by a binary outcome Roy model with imperfect foresight, identical to the binary outcome Roy model of Definition 1, except that the selection equation of Assumption 2 is replaced with the following:

Assumption 3. (*Imperfect foresight*) *The selection indicator satisfies $\mathbb{E}[Y_1 - Y_0 | \mathcal{I}] > 0 \Rightarrow D = 1$, $\mathbb{E}[Y_0 - Y_1 | \mathcal{I}] > 0 \Rightarrow D = 0$, where \mathcal{I} is the sigma-algebra characterizing the agent's information set at the time of sector choice.*

Definition 3 (Binary outcome Roy with imperfect foresight). *A model satisfying Assumptions 1 and 3, with $Y_0, Y_1 \in \{0, 1\}$, is called binary outcome Roy model with imperfect foresight.*

Our results in the next section characterize sharp bounds on the joint distributions of potential outcomes and highlight the difference in empirical content between perfect and imperfect foresight Roy models.

2.1. Sharp bounds for the binary outcome Roy model.

In the binary outcomes Roy model, the lack of point identification comes from the fact that the mapping from observed sector and success to unobserved skills is not single valued. We know that

when success in Sector 1 is observed, potential outcomes can be either $(Y_0 = 1, Y_1 = 1)$, i.e., success in both sectors, or $(Y_0 = 0, Y_1 = 1)$, i.e., success in Sector 1 only. Hence the identified probability of that a random individual in the population chooses Sector 1 and succeeds will not be sufficient to identify the probability of succeeding in Sector 1. What we do know, however, is that $Y = 0$ is observed if and only if the individual has neither the skills to succeed in Sector 0 nor in Sector 1. Hence, $\mathbb{P}(Y_0 = 0, Y_1 = 0) = \mathbb{P}(Y = 0)$. Moreover, if the individual has the skills to succeed in Sector 0, but not in Sector 1, then, success in Sector 0 will be observed, so that $\mathbb{P}(Y_0 = 1, Y_1 = 0) \leq \mathbb{P}(Y = 1, D = 0)$. Symmetrically, if the individual has the skills to succeed in Sector 1, but not in Sector 0, then, success in Sector 1 will be observed, so that $\mathbb{P}(Y_0 = 0, Y_1 = 1) \leq \mathbb{P}(Y = 1, D = 1)$.

The discussion above shows that the expressions hold. Showing sharpness of these bounds is more involved, and the proof of the Proposition 1 is given in the appendix, together with a more fastidious statement of the theorem, with a rigorous and unambiguous definition of *sharp bounds* in this context. Note that the *bounds* can take the form of an equality in case upper and lower bounds coincide.

Proposition 1 (Sharp bounds for the binary outcome Roy model). *The following equality and inequalities provide a set of sharp bounds for the joint distribution of potential outcomes (Y_0, Y_1) in the binary outcomes Roy model (Definition 1) and the alternative binary Roy model (Definition 2).*

$$\begin{aligned} \mathbb{P}(Y_0 = 1, Y_1 = 0) &\leq \mathbb{P}(Y = 1, D = 0), \\ \mathbb{P}(Y_0 = 0, Y_1 = 1) &\leq \mathbb{P}(Y = 1, D = 1), \\ \mathbb{P}(Y_0 = 0, Y_1 = 0) &= \mathbb{P}(Y = 0). \end{aligned} \tag{2.2}$$

The bounds in Proposition 1 summarize all the information in the (alternative) binary outcome Roy model about the joint distribution of potential outcomes. From these bounds, sharp bounds on the marginals, which are akin to traditional bounds on average treatment outcomes, can be recovered. Combining the equality and inequalities of (2.2), we obtain traditional bounds on the marginals (see for example Manski (2007), Section 7.5).

$$\mathbb{P}(Y = 1, D = 0) \leq \mathbb{E}Y_0 \leq \mathbb{P}(Y = 1) \text{ and } \mathbb{P}(Y = 1, D = 1) \leq \mathbb{E}Y_1 \leq \mathbb{P}(Y = 1). \tag{2.3}$$

If the means of marginal potential outcomes are the objects of interest, as in Manski (2007), Section 7.5, the bounds above are sharp without additional restrictions. Here, we take bounds on the joint distribution of potential outcomes as the object of interest. It is easy to see that (2.3) and $\mathbb{P}(Y_0 = 0, Y_1 = 0) = \mathbb{P}(Y = 0)$ are jointly equivalent to (2.2). However, from (2.3) alone, (2.2) cannot be recovered, so that information on the joint distribution is lost. The bounds on the average sector difference are

$$-\mathbb{P}(Y = 1, D = 0) \leq \mathbb{E}(Y_1 - Y_0) \leq \mathbb{P}(Y = 1, D = 1). \tag{2.4}$$

The sharp bounds of Proposition 1 emphasize two important facts:

- (1) On the one hand, despite the literature on *non identification* of competing risks, starting with Cox (1972) and Tsiatis (1975), the Roy model does in fact contain non trivial information about the joint distribution of potential outcomes, hence of skills, or more generally, of sector specific unobserved heterogeneity.
- (2) On the other hand, the sharp bounds of Proposition 1 can be very wide and they do not cross. For any joint distribution for (Y, D) , there exists a joint distribution for (Y_0, Y_1) that fits the binary outcome Roy model, so that the latter is not falsifiable in the absence of additional constraints.

Since the Roy model imposes strong restrictions on behaviour, the lack of testability is particularly vexing. We shall consider exclusion and monotonicity restrictions that allow us to recover testability of behaviour characterized by Roy sector selection. In the case of college major choice, considered in Section 4, one of our main concerns will be with explanations of the under representation of women in STEM. One candidate is wage discrimination in STEM, which is compatible with a Roy model of behaviour. Another is gender profiling in major choice, which is not. Hence the ability to test Roy maximizing behaviour in major selection is paramount.

2.2. Stochastically monotone instrumental variables (SMIV).

In order to allow falsifiability of the Roy model, we now investigate the implications of exclusion restrictions. Such exclusions are of two types: sector-specific variables, i.e., variables affecting only one outcome equation, but not the other (Assumption 6 in Appendix A.2), and variables that shift sector selection, but shift potential outcomes either not at all (Assumption 4 below), or only in one direction (Assumption 5 below). To sharpen the focus and save space, we discuss the conceptually relatively straightforward implications of sector specific variables in Appendix A, and consider mostly the effect of vectors Z of variables that affect sector selection, but have restricted impact on potential outcomes. We shall comment on the way in which sector specific exclusions modify the expressions and leave details to Appendix A.2 (As mentioned before, conditioning on remaining observed covariates is implicit in all the paper). We start the discussion with variables that shift selection, but not potential outcomes.

Assumption 4. *There exists a vector Z of observable random variables, such that $(Y_0, Y_1) \perp\!\!\!\perp Z$.*

Such variables are akin to typical instrumental variables, and examples within Roy models in the existing literature include parental education in Willis and Rosen (1979), distance to a college in Eisenhauer, Heckman, and Vytlacil (2015) and attendance in a Catholic high school in Altonjii, Elder, and Taber (2005). Local aggregate labour market variables at the time of sector selection are also often used, as in Eisenhauer, Heckman, and Vytlacil (2015) and references therein.

First, it is important to emphasize, that, unlike the generalized Roy model used in the contributions cited in the previous paragraph, the pure Roy selection mechanism imposes $D = 1$ when

$Y_1 > Y_0$ and $D = 0$ when $Y_1 < Y_0$. Hence, a selection shifter Z satisfying Assumption 4 can only affect the model in case of ties $Y_1 = Y_0$. The model is lexicographic, in the sense that agents care only about outcomes when choosing their sector of activity, unless the outcomes are equal in the two sectors, at which point other considerations guide their decision. As a result, Y is independent of Z , but (Y, D) is not jointly independent of Z , so that the bounds in Proposition 1 can be sharpened using variation in $\mathbb{P}(Y = 1, D = 1|Z)$ and in $\mathbb{P}(Y = 1, D = 0|Z)$. Taking the expressions in (2.2) conditionally on Z and using Assumption 4 to remove conditioning in the left-hand sides yields the bounds

$$\begin{aligned} \mathbb{P}(Y_0 = 1, Y_1 = 0) &\leq \inf_z \mathbb{P}(Y = 1, D = 0|Z = z), \\ \mathbb{P}(Y_0 = 0, Y_1 = 1) &\leq \inf_z \mathbb{P}(Y = 1, D = 1|Z = z), \\ \mathbb{P}(Y_0 = 0, Y_1 = 0) &= \mathbb{P}(Y = 0) = \mathbb{P}(Y = 0|Z = z). \end{aligned} \tag{2.5}$$

The tightened bounds are proven to be sharp in Appendix A and illustrated in Figure 2. The first two are *intersection bounds*, and inference can be carried out with the method proposed in Chernozhukov, Lee, and Rosen (2013).

The third expression in (2.5) gives a testable implication, since the binary outcomes Roy model under Assumptions 4 implies $Y \perp\!\!\!\perp Z$. We now argue that the latter summarizes all possible testable implications of the model. Indeed, for any joint distribution of (Y, D, Z) on $\{0, 1\}^2 \times \mathbb{R}$ satisfying $Y \perp\!\!\!\perp Z$, we can always define the pair of potential outcomes (Y_0, Y_1) by $Y_0 = Y_1 := Y$ and satisfy the constraints of the binary outcome Roy model under Assumption 4.

However, rejection of $Y \perp\!\!\!\perp Z$ cannot be attributed to a violation of the Roy selection assumption (Assumption 2) if the validity of the instrument is under question. In the case of college major choice, considered in Section 4, one of the proposed instrument is parental education. Unfortunately, the validity of this instrument is doubtful, as parental education level may be correlated with unobserved individual productivity in one or both of the sectors, hence affect potential outcomes directly. Indeed, Cunha, Heckman, and Schennach (2010) argue that cognitive and non-cognitive unobserved skills are determined in great part by parental environment and investment, which in turn is highly correlated with parental education. Distance to college is a similarly tainted instrument for returns to education, as discussed in Card (2001), since parental location preferences are correlated with unobserved cognitive and non cognitive parental investments. The same applies to local labor market conditions, which may drive endogenous location choices. Moreover, Kédagni and Mourifié (2016) derive sharp testable implications of $(Y_0, Y_1) \perp\!\!\!\perp Z$ within a binary potential outcomes model (without the Roy selection assumption) and their test tends to reject validity of parental education as an instrument, including in our data. More generally, instruments are elusive in the study of returns to education. The rest of this section is concerned with a weakening of Assumption 4 and a discussion of its validity, in order to recover testable implications of the Roy selection assumption.

Our objective now is to bring covariate information to bear and restore falsifiability of the Roy selection mechanism without relying on strong independence assumptions that are hard to substantiate. Although joint independence of potential labour market outcomes and parental education is indeed hard to substantiate, as unobserved benefits of parental education can raise productivity, it is natural to assume that increasing parental education cannot worsen potential labour market outcomes. Similarly, local aggregate labor market variables, such as the average wage in STEM for an individual socio-economic category at the time of college major decision, are also likely to be correlated with ex-post job market outcomes, but higher local average wages in STEM at the time of major decision are unlikely to produce lower wages in STEM at the time of graduation, barring complex general equilibrium adjustments.

The following weakening of Assumption 4 formalizes this insight, and applies to many selection shifters, that may not be valid instruments in the sense of Assumption 4, such as parental education, measures of school quality, merit based scholarships, distance to college or local labor market characteristics at the time of major selection. We adopt the following notion of monotonicity for the instrument. For details, refer to Shaked and Shanthikumar (2007), Section 6B. When comparing vectors, “ \geq ” denotes the componentwise partial order.

Definition 4. (*First Order Stochastic Dominance*) A distribution F_1 on \mathbb{R}^k is said to be first order stochastically dominated by a distribution F_2 if there exists random vectors X_1 with distribution F_1 and X_2 with distribution F_2 such that $\mathbb{P}(X_2 \geq X_1) = 1$. By extension, a random vector with distribution F_2 is also said to stochastically dominate a random vector with distribution F_1 .

Assumption 5. (*SMIV*) For any pair $z_2 \geq z_1$ in the support of a vector of observable variables Z , the conditional distribution of (Y_0, Y_1) given $Z = z_2$ first order stochastically dominates the distribution of (Y_0, Y_1) given $Z = z_1$ (denoted $(Y_0, Y_1)|Z = z_2 \succ_{FSD} (Y_0, Y_1)|Z = z_1$).

This monotonicity assumption is stronger than the monotone instrumental variable (MIV) of Manski and Pepper (2000), which requires only mean monotonicity of potential outcomes in the conditioning variable. SMIV provides sharper bounds than MIV on the joint distribution of potential outcomes, while being much easier to substantiate, in the case of parental education in particular, than the joint independence of Assumption 4. SMIV can hold with respect to a vector of instruments, which can increase the tightness of bounds on parameters of interest and the power of tests of the Roy selection mechanism derived below.

2.3. Sharp bounds and testability of the binary outcomes Roy model under SMIV.

An important distinction between the roles of the independence assumption (Assumption 4) and the stochastic monotonicity assumption (Assumption 5) is that, under the former, the instrument Z can only shift sector selection when $Y_0 = Y_1$, whereas under the latter, Z is no longer required

to be independent of potential outcomes (Y_0, Y_1) and can therefore induce variation in D , even when $Y_0 \neq Y_1$.

To see how the stochastic monotonicity assumption (Assumption 5) combines with the Roy selection mechanism (Assumption 2), start from the sharp bounds of Proposition 1 in the equivalent representation

$$\begin{aligned} \mathbb{P}(Y = 1, D = 0|Z = z) &\leq \mathbb{P}(Y_0 = 1|Z = z) \leq \mathbb{P}(Y = 1|Z = z), \\ \mathbb{P}(Y = 1, D = 1|Z = z) &\leq \mathbb{P}(Y_1 = 1|Z = z) \leq \mathbb{P}(Y = 1|Z = z), \\ \mathbb{P}(Y_0 = Y_1 = 0|Z = z) &= \mathbb{P}(Y = 0|Z = z). \end{aligned}$$

The statement $(Y_0, Y_1)|Z = z_2 \succsim_{FSD} (Y_0, Y_1)|Z = z_1$ is equivalent to $\mathbb{P}((Y_0, Y_1) \in U|Z = z_2) \geq \mathbb{P}((Y_0, Y_1) \in U|Z = z_1)$, for all upper sets U (Theorem 6.B.1 of Shaked and Shanthikumar (2007), Section 6B).

Definition 5. (*Upper Sets*) A subset U of a partially ordered set (\mathcal{S}, \geq) is called an upper set if $y \in U$ implies $\tilde{y} \in U$ for all $\tilde{y} \geq y$.

The non trivial upper subsets of $\{0, 1\}^2$ are

$$\{(1, 1)\}, \{(0, 1), (1, 1)\}, \{(1, 0), (1, 1)\}, \{(0, 1), (1, 0), (1, 1)\}.$$

Consider, for instance, the upper set $\{(0, 1), (1, 1)\}$. Stochastic monotonicity of (Y_0, Y_1) in z implies that $\mathbb{P}((Y_0, Y_1) \in \{(0, 1), (1, 1)\}|Z = z) \leq \mathbb{P}((Y_0, Y_1) \in \{(0, 1), (1, 1)\}|Z = \tilde{z})$ for all $\tilde{z} \geq z$, or equivalently $\mathbb{P}(Y_0 = 1|Z = z) \leq \mathbb{P}(Y_0 = 1|Z = \tilde{z})$. Since the latter is smaller than or equal to $\mathbb{P}(Y = 1|Z = \tilde{z})$ by Assumptions 1 and 2, we obtain $\mathbb{P}(Y_0 = 1|Z = z) \leq \mathbb{P}(Y = 1|Z = \tilde{z})$ for all $\tilde{z} \geq z$ in the domain of Z . Proceeding similarly with all upper subsets of $\{0, 1\}^2$, we obtain the following sharp bounds for the joint distribution of potential outcomes under Assumptions 1, 2 and 5: for all z in the domain of Z ,

$$\begin{aligned} \sup_{\tilde{z} \leq z} \mathbb{P}(Y = 1, D = 0|Z = \tilde{z}) &\leq \mathbb{P}(Y_0 = 1|Z = z) \leq \inf_{\tilde{z} \geq z} \mathbb{P}(Y = 1|Z = \tilde{z}), \\ \sup_{\tilde{z} \leq z} \mathbb{P}(Y = 1, D = 1|Z = \tilde{z}) &\leq \mathbb{P}(Y_1 = 1|Z = z) \leq \inf_{\tilde{z} \geq z} \mathbb{P}(Y = 1|Z = \tilde{z}), \\ \sup_{\tilde{z} \geq z} \mathbb{P}(Y = 0|Z = \tilde{z}) &\leq \mathbb{P}(Y_0 = Y_1 = 0|Z = z) \leq \inf_{\tilde{z} \leq z} \mathbb{P}(Y = 0|Z = \tilde{z}). \end{aligned} \tag{2.6}$$

The third line of the display in (2.6) combines identification of $\mathbb{P}(Y_0 = Y_1 = 0|Z = z)$, which is equal to $\mathbb{P}(Y = 0|Z = z)$, and the testable implications $\sup_{\tilde{z} \geq z} \mathbb{P}(Y = 0|Z = \tilde{z}) \leq \inf_{\tilde{z} \leq z} \mathbb{P}(Y = 0|Z = \tilde{z})$ for all z in the domain of Z . The latter is equivalent to stochastic monotonicity of Y in z , which turns out to summarize all testable implications of Roy under Assumption 5 as formalized in the following theorem.

Theorem 1 (Sharp bounds and testable implications of Roy under SMIV).

- (1) The display in (2.6) characterizes the identified set for the joint distribution of potential outcomes in the binary outcomes Roy model under Assumptions 1, 2 and 5.
- (2) Under Assumptions 1, 2 and 5, the following holds: (*) For any pair $z_2 \geq z_1$ in the support of the vector of observable variables Z , $Y|Z = z_2 \succsim_{FSD} Y|Z = z_1$.

(3) If (Y, Z) satisfies $(*)$, then there is a pair (Y_0, Y_1) such that Assumptions 1, 2 and 5 hold.

When Y is stochastically monotone in z , $\inf_{\tilde{z} \geq z} \mathbb{P}(Y = 1|Z = \tilde{z})$ is equal to $\mathbb{P}(Y = 1|Z = z)$, which, by the third line of (2.6) is equal to $\mathbb{P}(Y_0 = Y_1 = 0|Z = z)$. Hence the right-hand side inequalities in the first two lines of (2.6) are redundant, and the identified set for the joint distribution of potential outcomes is characterized by two inequalities and one equality. The proof of Theorem 1(1) is given in Appendix A and the identified set is represented graphically on the right-hand-side panel of Figure 2. The proof of Theorem 1(2,3) is straightforward. Indeed, under Assumptions 1 and 2, we have $\mathbb{P}(Y_0 \leq y, Y_1 \leq y|Z) = \mathbb{P}(Y \leq y|Z)$, for all y , since counterfactual outcomes cannot be larger than realized ones. Hence stochastic monotonicity of (Y_0, Y_1) immediately implies stochastic monotonicity of Y . We now argue that it constitutes a sharp testable implication of the Roy selection mechanism. Indeed, given any joint distribution of observable variables (Y, D, Z) on $\mathbb{R} \times \{0, 1\} \times \mathbb{R}^d$, with $Y|Z = z_2 \succsim_{FSD} Y|Z = z_1$ for each $z_2 \geq z_1$, the pair of potential outcomes (Y_0, Y_1) can always be chosen in such a way that Assumptions 1, 2 and 5 are satisfied. For example, setting $Y_0 = Y_1 = Y$ would satisfy all the constraints.

Theorem 1 shows that testing the Roy selection mechanism simply boils down to testing stochastic monotonicity of observed outcomes with respect to the monotone instrumental variable, which can be performed with existing inference methods in Lee, Linton, and Whang (2009), Delgado and Escanciano (2012) and Hsu, Liu, and Shi (2016). Statements (2) and (3) of Theorem 1 make no mention of the binary outcomes Roy model, since they are valid without restrictions on the domain of the outcome variables.

2.4. Imperfect foresight.

To address the concern that rejection of the Roy selection mechanism may be down to rejecting the assumption that agents are perfectly informed of their future potential outcomes at the time of sector selection, we also derive testable implications of the binary outcome Roy model with imperfect foresight of Definition 3. The latter is identical to the binary outcome Roy model, except that the Roy selection assumption, Assumption 2, is replaced by imperfect foresight, namely Assumption 3.

Under the potential outcomes model, i.e., Assumption 1, only, we still know that an individual with the skill to succeed in Sector 0, but not in Sector 1, will be observed as having succeeded in Sector 0 or as having failed in Sector 1 (the latter was ruled out under the Roy selection rule of Assumption 2). Hence $\mathbb{P}(Y_0 = 1, Y_1 = 0|Z) \leq \mathbb{P}(Y = 1, D = 0|Z) + \mathbb{P}(Y = 0, D = 1|Z)$ and symmetrically for $\mathbb{P}(Y_0 = 0, Y_1 = 1|Z)$. An individual without the skills to succeed in either sector will be observed to fail, so that $\mathbb{P}(Y_0 = 0, Y_1 = 0|Z = z) \leq \mathbb{P}(Y = 0|Z = z)$. Under stochastic monotonicity in Z (Assumption 5), the latter yields $\mathbb{P}(Y_0 = 0, Y_1 = 0|Z = z) \leq \inf_{\tilde{z} \leq z} \mathbb{P}(Y = 0|Z = \tilde{z})$. In addition, observing success in Sector d necessary implies that the agent has the skills required

for Sector d , hence $\mathbb{P}(Y_d = 1|Z) \geq \mathbb{P}(Y = 1, D = d|Z)$. Under Assumption 5, the latter yields $\mathbb{P}(Y_d = 1|Z = z) \geq \sup_{\tilde{z} \leq z} \mathbb{P}(Y = 1, D = d|Z = \tilde{z})$, $d \in \{0, 1\}$.

We now add selection information according to Assumption 3. The latter is equivalent to

$$Y = Y_d \Rightarrow \mathbb{E}[Y|\mathcal{I}] = \mathbb{E}[Y_d|\mathcal{I}] \geq \mathbb{E}[Y_{1-d}|\mathcal{I}], \quad d = 0, 1.$$

After integration, assuming that $\mathcal{I} = \sigma(Z)$, this yields $\mathbb{E}(Y|Z) = \max\{\mathbb{E}(Y_0|Z), \mathbb{E}(Y_1|Z)\}$. Under Assumption 5, the latter yields monotonicity of $\mathbb{E}(Y|Z = z)$ in z . Putting it all together yields the following sharp bounds on the joint distribution of potential outcomes under Assumptions 1, 3 and 5:

$$\begin{aligned} \mathbb{P}(Y_0 = 1, Y_1 = 0|Z = z) &\leq \mathbb{P}(Y = 1, D = 0|Z = z) + \mathbb{P}(Y = 0, D = 1|Z = z), \\ \mathbb{P}(Y_0 = 0, Y_1 = 1|Z = z) &\leq \mathbb{P}(Y = 0, D = 0|Z = z) + \mathbb{P}(Y = 1, D = 1|Z = z), \\ \mathbb{P}(Y_0 = 0, Y_1 = 0|Z = z) &\leq 1 - \mathbb{E}(Y|Z = z), \end{aligned} \quad (2.7)$$

and

$$\begin{aligned} \sup_{\tilde{z} \leq z} \mathbb{P}(Y = 1, D = 0|Z = \tilde{z}) &\leq \mathbb{E}(Y_0|Z = z), \\ \sup_{\tilde{z} \leq z} \mathbb{P}(Y = 1, D = 1|Z = \tilde{z}) &\leq \mathbb{E}(Y_1|Z = z), \\ \max\{\mathbb{E}(Y_0|Z = z), \mathbb{E}(Y_1|Z = z)\} &= \mathbb{E}[Y|Z = z], \end{aligned} \quad (2.8)$$

for all z in the support of Z .

The inequalities above define the identified set for the joint distribution of potential outcomes. Testable implications of the Roy model with imperfect foresight include monotonicity of $\mathbb{E}[Y|Z = z]$ in z as derived above, which proves Theorem 2(2) below. It can be easily shown that in the binary case, this monotonicity summarizes the empirical content of the Roy selection assumption with imperfect foresight, as stated in Theorem 2. Indeed, for any given vector (Y, D, Z) such that $\mathbb{E}[Y|Z = z]$ is non decreasing in z , setting $Y_0 = Y_1 = Y$ satisfies the assumptions, which proves Theorem 2(3) below.

Theorem 2 (Testable implications of Roy with imperfect foresight).

- (1) *The displays in (2.7) and (2.8) jointly characterize the identified set for the joint distribution of potential outcomes in the binary outcomes Roy model with imperfect foresight under Assumptions 1, 3 and 5 with $\mathcal{I} = \sigma(Z)$.*
- (2) *If Assumptions 1, 3, 5 hold with \mathcal{I} -measurable Z , then $\mathbb{E}[Y|Z = z]$ is non decreasing in z .*
- (3) *For any distribution G on $\{0, 1\}^2 \times \text{Supp}(Z)$, such that $\mathbb{E}[Y|Z = z]$ is non decreasing in z , there exists a random vector $(Y_0, Y_1, D, Z) \in \{0, 1\}^3 \times \text{Supp}(Z)$ such that $(Y_1 D + Y_0(1 - D), D, Z)$ has distribution G and Assumptions 3 and 5 are satisfied with $\mathcal{I} = \sigma(Z)$.*

We can therefore test Roy with imperfect foresight under Assumption 5 simply by testing monotonicity of $\mathbb{E}[Y|Z = z]$ in z , using existing inference methods in Chetverikov (2013) or Hsu, Liu, and Shi (2016). We can also verify that stochastic monotonicity of Y (the testable implication of Roy selection as shown in Theorem 1) does indeed imply monotonicity of $\mathbb{E}[Y|Z = z]$ in z , which

is consistent with the fact that Assumption 2 implies Assumption 3. Moreover, in the binary outcomes case, the testable implications of Roy behavior under SMIV and those of imperfect foresight Roy under SMIV are identical, since when Y is binary, stochastic monotonicity of $Y|Z = z$ and monotonicity of $\mathbb{E}[Y|Z = z]$ are equivalent. Hence, rejection of Roy selection behavior under SMIV implies rejection of Roy with imperfect foresight as well. However, the identified set for the joint distribution of potential outcomes in Theorem 1(1) is nested in, and weakly tighter than the identified set of Theorem 2(1), since the combination of Assumptions 2 and 5 contains more information on the joint distribution of potential outcomes than the combination of Assumptions 3 and 5.

2.5. Bounds on departures from Roy selection.

In case of rejection of the Roy selection mechanism, the methodology developed here, and particularly the information on the joint distribution of potential outcomes, allows us to quantify departures from the Roy sector selection rule (Assumption 2). If agents are believed to be expected outcome maximizers, i.e., to behave according to the binary outcome Roy model with imperfect foresight, this measure of departure can be interpreted as a measure of the cost of imperfect foresight. If, on the other hand, departures from the Roy model with imperfect foresight are entertained, then the measure of departure we propose also captures the extent to which considerations other than potential outcome maximization enter in the decision. This may be the result of maximization of a utility function that depends on aspects beyond the chosen outcomes. It may also reveal a bias in decision making. This would be the case, in our application to major choice, if gender profiling discouraged women from choosing STEM majors.

Departure from Roy sector selection, which we therefore interpret as inefficiency of sector choice, can be measured as the difference between maximum potential outcome and realized outcome, the two being equal by definition in the case of Roy selection according to Assumption 2.

Definition 6 (Efficiency loss). *Efficiency loss from Roy selection departures is defined for each $z \in \text{Supp}(Z)$ as $el(z) := \mathbb{P}(\max(Y_0, Y_1) = 1|Z = z) - \mathbb{P}(Y = 1|Z = z)$ in the binary outcomes case, and, for each $y \in \text{Supp}(Y)$, as $el(y, z) := \mathbb{P}(Y \leq y|Z = z) - \mathbb{P}(\max(Y_0, Y_1) \leq y|Z = z)$, otherwise.*

We have $\mathbb{P}(\max(Y_0, Y_1) = 1|Z = z) - \mathbb{P}(Y = 1|Z = z) = \mathbb{P}(Y = 0|Z = z) - \mathbb{P}(Y_0 = Y_1 = 0|Z = z)$. Since in the binary outcomes Roy model, $\mathbb{P}(Y_0 = Y_1 = 0|Z = z)$ is identified as $\mathbb{P}(Y = 0|Z = z)$, efficiency loss is zero, which justifies the interpretation as a departure from Roy selection (Assumption 2). When Assumption 2 is dropped, efficiency loss is non negative. Since $\mathbb{P}(Y = 0|Z = z)$ is identified, bounds on efficiency loss or departure from Roy will be obtained from bounds on $\mathbb{P}(Y_0 = Y_1 = 0|Z = z)$ under Assumptions 1 and 5 only. Since $\mathbb{P}(Y_0 = Y_1 = 0|Z = z)$ involves the joint distribution of potential outcomes, sharp bounds on marginal distributions alone cannot deliver the desired bounds on efficiency loss. This feature is shared by other policy relevant parameters such as ex-post regret, skill correlation, sector effect conditional on the chosen sector

and the proportion who benefit from a given sector, all of which can also be bound using this methodology.

Under Assumptions 1 and 5, the identified set for the joint distribution of potential outcomes is obtained in a similar fashion to (2.7)-(2.8), except that we cannot rely on selection information, so that the upper bounds in (2.8) are obtained from the fact that talent for Sector d only precludes observing failure in Sector d . Bounds (2.7)-(2.8) are therefore replaced by

$$\begin{aligned} \mathbb{P}(Y_0 = 1, Y_1 = 1|Z = z) &\leq \inf_{\tilde{z} \geq z} \mathbb{P}(Y = 1|Z = \tilde{z}) \\ \mathbb{P}(Y_0 = 0, Y_1 = 0|Z = z) &\leq \inf_{\tilde{z} \leq z} \mathbb{P}(Y = 0|Z = \tilde{z}) \\ \mathbb{P}(Y_0 = 1, Y_1 = 0|Z = z) &\leq \mathbb{P}(Y = 1, D = 0|Z = z) + \mathbb{P}(Y = 0, D = 1|Z = z) \\ \mathbb{P}(Y_0 = 0, Y_1 = 1|Z = z) &\leq \mathbb{P}(Y = 0, D = 0|Z = z) + \mathbb{P}(Y = 1, D = 1|Z = z), \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} \sup_{\tilde{z} \leq z} \mathbb{P}(Y = 1, D = 0|Z = \tilde{z}) &\leq \mathbb{P}(Y_0 = 1|Z = z) \leq 1 - \sup_{\tilde{z} \geq z} \mathbb{P}(Y = 0, D = 0|Z = \tilde{z}) \\ \sup_{\tilde{z} \leq z} \mathbb{P}(Y = 1, D = 1|Z = \tilde{z}) &\leq \mathbb{P}(Y_1 = 1|Z = z) \leq 1 - \sup_{\tilde{z} \geq z} \mathbb{P}(Y = 0, D = 1|Z = \tilde{z}), \end{aligned} \tag{2.10}$$

for all z in the support of Z . From (2.9), we obtain immediately an upper bound on $\mathbb{P}(Y_0 = Y_1 = 0|z)$, namely $\inf_{\tilde{z} \leq z} \mathbb{P}(Y = 0|Z = \tilde{z})$. Sharp bounds are obtained by projecting (2.9)-(2.10) onto component $\mathbb{P}(Y_0 = Y_1 = 0|z)$ in the 3-simplex.

Proposition 2 (Bounds on efficiency loss).

- (1) *The displays in (2.9) and (2.10) jointly characterize the identified set for the joint distribution of potential outcomes under Assumptions 1 and 5, with $Y \in \{0, 1\}$.*
- (2) *Under Assumptions 1 and 5 with $Y \in \{0, 1\}$, efficiency loss due to departures from Roy selection satisfies, for each $z \in \text{Supp}(Z)$, $el(z) \geq \mathbb{P}(Y = 0|Z = z) - \inf_{\tilde{z} \leq z} \mathbb{P}(Y = 0|Z = \tilde{z})$.*

The results on efficiency loss allow us to identify values of observable characteristics that induce the most costly misallocation of talent and field of study and are therefore prime targets for intervention.

3. ROY MODEL WITH DISCRETE-CONTINUOUS OUTCOMES

Extending the analysis to richer sets of outcomes, including mixed discrete and continuous potential outcomes does not remove the lack of identification issue in the Roy model (and the related competing risks model). The range of observables is richer, but so is the object of interest, i.e., the joint distribution of potential outcomes. Given partial observability and endogenous sector selection, the Roy model is essentially partially identified. Results obtained in the form of sharp bounds on the joint distribution of potential outcomes and the methods used to derive them are analogous to the corresponding results and methods in the binary outcome case, except in one important respect. When considering distributional aspects, such as inequality, the distinction between *pointwise bounds* and *functional bounds* is crucial as described below. After a discussion of the latter point, we

proceed to analyze testability and quantifying departures from the Roy selection mechanism along the same lines as in the binary outcomes case in Section 2.

3.1. Functionally sharp bounds for the Roy model.

Consider the Roy model of Section 1, under Assumptions 1 and 2. Bounds on the marginal distributions of potential outcomes can be derived very easily as follows. For any real number y , $\mathbb{P}(Y_d \leq y) = \mathbb{P}(Y_d \leq y, D = d) + \mathbb{P}(Y_d \leq y, D = 1 - d)$. The first term on the right-hand-side is identified. The second term on the right-hand-side is bounded below by $\mathbb{P}(Y_{1-d} \leq y, D = 1 - d)$, and above by $\mathbb{P}(D = 1 - d)$. The resulting bounds were shown by Peterson (1976) to be *pointwise sharp* for the marginal distributions of potential outcomes, in the sense that any pair of distributions of potential outcomes that satisfy the bounds for a given fixed y , can be obtained from some joint distribution of observable variables (Y, D) under the assumptions of the Roy model. However, as Crowder (1991) pointed out, there are additional non redundant cross quantile restrictions, namely, for all $y \geq x$, $\mathbb{P}(x < Y_d \leq y) \geq \mathbb{P}(x < Y_d \leq y, D = d)$. If the object of interest involves densities, such as the hazard rate, or functional features, such as inequality measures, the difference between the latter bounds and pointwise bounds can be considerable. Indeed, combining Peterson bounds involves an additional term $-\mathbb{P}(D = 1 - d)$ in the lower bound. This difference arises because the monotonicity of the distribution function is not factored in. Graphically, the difference between pointwise bounds and functional bounds can be highlighted on Figure 1. A candidate distribution function for Y_d that is drawn through the two points (\tilde{y}_1, q_1) and (\tilde{y}_2, q_2) can lie between the curves $\mathbb{P}(Y \leq y)$ and $\mathbb{P}(Y \leq y, D = d) + \mathbb{P}(D = 1 - d)$. Hence it satisfies pointwise bounds. However, its slope is lower in some regions than the slope of the curve $\mathbb{P}(Y \leq y, D = d)$, so that it fails to satisfy the functional bounds.

Turning to the joint distribution function of potential outcomes, pointwise bounds can also be derived very easily. Indeed, we immediately have

$$\mathbb{P}(Y \leq \min(y_0, y_1)) \leq \mathbb{P}(Y_0 \leq y_0, Y_1 \leq y_1) \leq \mathbb{P}(Y \leq y_0, D = 0) + \mathbb{P}(Y \leq y_1, D = 1). \quad (3.1)$$

Corollary 1 of Bedford and Meilijson (1997) shows that the bounds (3.1) can be attained under their competing risks specification. However, once again, these bounds fail to incorporate monotonicity conditions, and they can entail loss of information, when describing functional features of potential outcomes.

The object of interest is the joint distribution (Y_0, Y_1) , the information on which we wish to characterize using the identified joint distribution of observable variables (Y, D) . Take any subset A of \mathbb{R}^2 and consider bounding the probability of $(Y_0, Y_1) \in A$. If A contains points (y_0, y) such that $y_0 \leq y$ can give rise to observation $(Y = y, D = 1)$, and if A contains points (y, y_1) such that $y_1 \leq y$ can give rise to observation $(Y = y, D = 0)$. Hence, observation $(Y = y, D = d)$ such that $y \in U_{A,0}$ below, and only those, can be rationalized by elements of A . Similarly, to derive the lower bound,

notice that $(Y, D) = (y, 1)$ can arise for any $(Y_0, Y_1) \in [\underline{b}, y] \times \{y\}$, so that (Y_0, Y_1) mass could be concentrated outside A unless the whole of $[\underline{b}, y] \times \{y\}$ is contained in A (where \underline{b} is common lower bound of the supports of Y_0 and Y_1).

Definition 7. For any Borel set A in \mathbb{R}^2 , define the sets $U_{A,0}$, $U_{A,1}$ and $L_{A,0}$, $L_{A,1}$ as

$$\begin{aligned} U_{A,0} &= \{y \in \mathbb{R} \mid \{y\} \times [\underline{b}, y] \cap A \neq \emptyset\}, & L_{A,0} &= \{y \in \mathbb{R} \mid \{y\} \times [\underline{b}, y] \subseteq A\}, \\ U_{A,1} &= \{y \in \mathbb{R} \mid [\underline{b}, y] \times \{y\} \cap A \neq \emptyset\}, & L_{A,1} &= \{y \in \mathbb{R} \mid [\underline{b}, y] \times \{y\} \subseteq A\}. \end{aligned}$$

We shall formally show that the upper bound is $\mathbb{P}(Y \in U_{A,0}, D = 0) + \mathbb{P}(Y \in U_{A,1}, D = 1)$. Similarly, the lower bound will be shown to be $\mathbb{P}(Y \in L_{A,0}, D = 0) + \mathbb{P}(Y \in L_{A,1}, D = 1)$. In the case, where A is an upper set (Definition 5), the bounding sets of Definition 7 take a very simple form, and we derive the constraints associated with Assumption 5 (SMIV) accordingly.

Theorem 3 (Sharp bounds for the joint distribution).

- (1) Let the distribution of observable variables (Y, D) on $\mathbb{R} \times \{0, 1\}$ be given. Under Assumptions 1 and 2, the distribution of (Y_0, Y_1) on \mathbb{R}^2 satisfies, for all Borel subset A of \mathbb{R}^2 ,

$$\begin{aligned} \mathbb{P}(Y \in L_{A,0}, D = 0) + \mathbb{P}(Y \in L_{A,1}, D = 1) \\ \leq \mathbb{P}((Y_0, Y_1) \in A) \\ \leq \mathbb{P}(Y \in U_{A,0}, D = 0) + \mathbb{P}(Y \in U_{A,1}, D = 1). \end{aligned}$$

- (2) Conversely, for any joint distribution satisfying the bounds above, there exists a pair (Y_0, Y_1) with that distribution, which satisfies Assumptions 1 and 2.
- (3) If Assumption 5 also holds, then the distribution of (Y_0, Y_1) also satisfies, for all upper set A of \mathbb{R}^2 , all $z \in \text{Supp}(Z)$,

$$\begin{aligned} \sup_{\tilde{z} \leq z} [\mathbb{P}(Y \geq \underline{y}_0^A, D = 0 \mid Z = \tilde{z}) + \mathbb{P}(Y \geq \underline{y}_1^A, D = 1 \mid Z = \tilde{z})] \\ \leq \mathbb{P}((Y_0, Y_1) \in A \mid Z = z) \leq \inf_{\tilde{z} \geq z} \mathbb{P}(Y \geq \underline{y}^A \mid Z = \tilde{z}), \end{aligned}$$

with $\underline{y}^A := \inf\{y : (y, y) \in A\}$, $\underline{y}_0^A := \inf\{y : (y, +\infty) \times \mathbb{R} \subseteq A\}$ and $\underline{y}_1^A := \inf\{y : \mathbb{R} \times (y, +\infty) \subseteq A\}$.

Theorem 3(1) allows us to easily recover Peterson bounds with suitable choices of A . Choosing $A = [\underline{b}, y] \times \mathbb{R}$ yields Peterson bounds on the marginal distribution of Y_0 . Choosing $A = [\underline{b}, y_0] \times [\underline{b}, y_1]$ yields Peterson bounds on the joint distribution of (Y_0, Y_1) . Finally, applying Theorem 3 to sets of the form $(y_1, y_2] \times \mathbb{R}$ and $\mathbb{R} \times (y_1, y_2]$ yields the following bounds on the marginal distributions of Y_d , for $d = 1, 0$:

$$\mathbb{P}(y_1 < Y \leq y_2, D = d) + \mathbb{P}(Y \leq y_2, D = 1 - d)1\{y_1 = \underline{b}\} \tag{3.2}$$

$$\leq \mathbb{P}(y_1 < Y_d \leq y_2)$$

$$\leq \mathbb{P}(y_1 < Y \leq y_2, D = d) + \mathbb{P}(y_1 < Y, D = 1 - d). \tag{3.3}$$

The upper bound (3.3) is redundant. Indeed, it can be recovered from lower bounds on $\mathbb{P}(y_2 < Y_d \leq \infty)$ and $\mathbb{P}(\underline{b} < Y_d \leq y_1)$. We shall show that the class of sets of the form $(y_1, y_2] \times \mathbb{R}$ and

$\mathbb{R} \times (y_1, y_2]$ suffice to characterize the marginal potential distributions and that the lower bounds are functionally sharp, as formulated in Corollary 1 below. The bounds are similar, though not identical, to the bounds in Theorem 1 of Bedford and Meilijson (1997) for a related competing risks model, discussed in the paragraph below Assumption 2. The result is proved in the appendix, with a more rigorous statement and formal definition of *functional sharp bounds*.

Corollary 1 (Sharp bounds for the marginal distributions). *Under Assumptions 1 and 2, the bounds*

$$\mathbb{P}(y_1 < Y_d \leq y_2) \geq \mathbb{P}(y_1 < Y \leq y_2, D = d) + \mathbb{P}(Y \leq y_2, D = 1 - d)1\{y_1 = \underline{b}\}$$

for all $y_1, y_2 \in \mathbb{R} \cup \{\pm\infty\}$, $y_1 < y_2$, and $d = 0, 1$, are *functional sharp bounds*.

Corollary 1 tells us that intervals are sufficient to characterize all the information we have on the marginal distribution of potential outcomes (Y_1, Y_0) . They form a *core determining* class of sets, in the terminology of Galichon and Henry (2006, 2011). This has several advantages. It allows the incorporation of exclusion restrictions and lends itself to the partial identified inference of Chernozhukov, Lee, and Rosen (2013) and Andrews and Shi (2014). The characterization of Corollary 1 allows us to derive sharp bounds on functional features such as measures of inequality.

3.2. Testing Roy and bounding departures from Roy.

As in the binary outcome case of Section 2, the Roy model defined by Assumptions 1 and 2, is not falsifiable without additional information. Indeed, for any joint distribution (Y, D) , potential outcomes (Y_0, Y_1) can be chosen, for instance with $Y_0 = Y_1 = Y$, such that Assumptions 1 and 2 hold. Given the unavailability of an instrument that satisfies Assumption 4, we examine falsifiability of the model under the stochastically monotone instrumental variable assumption (Assumption 5). Theorem 1(2,3) shows that stochastic monotonicity of observed outcomes with respect to the instrument summarizes all observable implications of the Roy selection mechanism under SMIV. Hence Roy selection behavior can be tested using existing inferential methods to test stochastic monotonicity. As concerns falsifiability of the Roy model with imperfect foresight, Theorem 2(2) shows that a testable implication is monotonicity of $\mathbb{E}[Y|Z = z]$ in z , which can also be tested using existing inference methods on regression monotonicity. However, Theorem 2(3) only holds in the binary outcomes case, since monotonicity of $\mathbb{E}[Y|Z = z]$ does not otherwise imply stochastic monotonicity of $Y|Z = z$ in z , and therefore does not summarize the empirical content of Roy with imperfect foresight under Assumption 5.

According to Definition 6, departure from Roy selection behavior or inefficiency of sector choice can be measured with the difference $\mathbb{P}(Y \leq y|Z = z) - \mathbb{P}(\max(Y_0, Y_1) \leq y|Z = z)$. The latter is zero under Assumption 2 (Roy selection mechanism). Otherwise, $\mathbb{P}(Y \leq y|Z = z) - \mathbb{P}(\max(Y_0, Y_1) \leq y|Z = z) \geq \mathbb{P}(Y \leq y|Z = z) - \inf_{\tilde{z} \leq z} \mathbb{P}(Y \leq y|Z = \tilde{z})$, under Assumption 5.

Proposition 3. *Under Assumptions 1 and 5, efficiency loss of Definition 6 satisfies*

$$el(y, z) \geq \mathbb{P}(Y \leq y | Z = z) - \inf_{\tilde{z} \leq z} \mathbb{P}(Y \leq y | Z = \tilde{z}),$$

for all (y, z) in the support of (Y, Z) .

As the binary case, the lower bound on efficiency loss is zero under Assumption 2 and can serve to construct a test statistic for a test of Roy selection behavior.

3.3. Functional features of potential distributions.

The original motivation of the Roy model was to analyze the effect of self-selection on wage distributions, and particularly on wage inequality. Heckman and Honoré (1990) show that self-selection reduces aggregate inequality when skills are log normal and within sector inequality when skills have a log concave distribution. One of the purposes of functional sharp bounds derived in the previous section is to analyze the effect of self-selection on inequality of potential outcomes in the specification of the Roy model we consider here, where the Roy model structure is stripped down to the self-selection mechanism. Functional sharp bounds on the potential outcome distributions allow us to derive sharp bounds on inequality measures. In this section, we concentrate on the interquantile range, although the same reasoning applies to other functionals from the vast literature on distributional inequality.

Consider two quantiles q_1 and q_2 with $q_2 > q_1$, as illustrated on Figure 1. The most commonly used range is the interquartile range, where $q_1 = 1 - q_2 = 1/4$, but other cases, such as $q_1 = 1 - q_2 = 0.1$, are also of great empirical relevance. Peterson bounds on the distribution of Y_d impose $\mathbb{P}(Y_d \leq y_1) \leq \mathbb{P}(Y \leq y_1, D = d) + \mathbb{P}(D = 1 - d) = q_1$ and $\mathbb{P}(Y_d \leq y_2) \geq \mathbb{P}(Y \leq y_2) = q_2$. Hence, the upper bound on the interquantile range based on pointwise sharp bounds for the distribution of Y_d is $y_2 - y_1$. However, functional sharp bounds of Corollary 1 are violated, since $q_2 - q_1 < \mathbb{P}(y_1 < Y \leq y_2, D = d)$. On Figure 4, we exhibit another pair of points, namely (\tilde{y}_1, q_1) and (\tilde{y}_2, q_2) such that a distribution for potential Y_d cannot cross these two points and satisfy the functional sharp bounds of Corollary 1.

We now show how to derive sharp bounds for the interquantile range. For ease of notation throughout this section, for $d = 0, 1$, and for each $y \in \mathbb{R}$, denote $F(y) := \mathbb{P}(Y \leq y)$, $F_d(y) := \mathbb{P}(Y_d \leq y)$, $\underline{F}_d(y) := \mathbb{P}(Y \leq y, D = d)$, $\bar{F}_d(y) := \mathbb{P}(Y \leq y, D = d) + \mathbb{P}(D = 1 - d)$, and f^{-1} the generalized inverse of f , i.e., $f^{-1}(q) = \inf\{y : f(y) > q\}$. Start from any y within the pointwise quantile bounds $\bar{F}_d^{-1}(q_1) \leq y \leq F^{-1}(q_2)$. From y , the largest interquantile range obtains in either of the following two cases:

- (1) when $F(y)$ is hit first, in which case the interquantile range is $F^{-1}(q_2) - y$,

the *self-selection economy*, to distributions of outcomes that would result from random assignment of individuals to sectors of activity, hereafter called outcome distributions in the *random assignment economy*, as in Heckman and Sedlacek (1985),(1990) and Heckman and Honoré (1990). In Sector d , the distribution of outcomes in the random assignment economy is the distribution of potential outcome Y_d , while the distribution of outcomes of the self-selection economy is $\mathbb{P}(Y \leq y|D = d)$. In the aggregate population, the distribution of outcomes of the random assignment economy is $\mathbb{P}(Y_0 \leq y)\mathbb{P}(D = 0) + \mathbb{P}(Y_1 \leq y)\mathbb{P}(D = 1)$, whereas the distribution of outcomes of the self-selection economy is simply the distribution of observable outcomes Y . These cases are collected in Table 1.

In Sector d , the interquantile range between quantiles q_1 and q_2 of the distribution of outcomes in the random assignment economy is bounded above by (3.4). In the self-selection economy, it is identified as the interquantile range of the distribution $\mathbb{P}(Y \leq y|D = d)$. The following proposition shows how they compare.

Proposition 4 (Inequality in Sector d).

- (1) *If the distribution of outcomes Y conditional on $D = d$ first order stochastically dominates the distribution of outcomes Y conditional on $D = 1 - d$, i.e., $\mathbb{P}(Y \leq y|D = d) \leq \mathbb{P}(Y \leq y|D = 1 - d)$ for all $y \in \mathbb{R}$, then, for any pair of quantiles, the interquantile range of the distribution of outcomes in Sector d in the self-selected economy is lower than the upper bound of the interquantile range of the distribution of outcomes in Sector d in the random assignment economy.*
- (2) *If the stochastic dominance relation of (1) does not hold, then there exists distributions for the pair (Y, D) such that the interquantile range of the distribution of outcomes in Sector d in the self-selected economy is larger than the upper bound of the interquantile range of the distribution of outcomes in Sector d in the random assignment economy.*

Proposition 4 tells us two things. On the one hand, if Sector d is unambiguously more profitable in the self-selected economy, it is possible for inequality in Sector d , as measured by the interquantile range, to decrease with self-selection, relative to an economy with random assignment of individuals to sectors. On the other hand, if neither sector dominates the other in the self-selection economy, then there are joint distributions of observables under which we know that self-selection unambiguously increases inequality in Sector d . In case no sector stochastically dominates the other, the hypothesis that self-selection increases inequality is testable based on the bounds of (3.4), in the sense that one

Table 1. Distribution of outcomes under self-selection and random assignment.

	self-selection economy	random economy
Sector d	$\mathbb{P}(Y \leq y D = d)$	$\mathbb{P}(Y_d \leq y)$
Aggregate	$\mathbb{P}(Y \leq y)$	$\mathbb{P}(Y_0 \leq y)\mathbb{P}(D = 0) + \mathbb{P}(Y_1 \leq y)\mathbb{P}(D = 1)$

can test the hypothesis that the interquantile range in the self-selected economy is larger than the upper bound of the interquantile range in the randomized economy.

4. ROY MODEL OF COLLEGE MAJOR CHOICE IN CANADA AND GERMANY

Since James, Alsalam, Conaty, and To (1989) pointed out that major choice mattered more to labour market outcomes than college choice, the literature on returns to college education has placed some focus on the determinants of major choice and the effects on labour market outcomes. The salient classification that has come to dominate the debate is between STEM and non-STEM degrees, and there is ample evidence of the labour market advantages conferred on male graduates by STEM degrees: James, Alsalam, Conaty, and To (1989) and Arcidiacono (2004) for the US, Kelly, O’Connell, and Smyth (2010) for Ireland, Chevalier (2011) for the UK, Maselli and Beblavý (2014) for several EU countries.

The wage benefits of STEM degrees have been found to be a significant but not sole determinant of major choice. Arcidiacono (2004) finds that high ability students view education as a consumption good in the US. Beffy, Fougère, and Maurel (2012) find elasticity of major choice to expected income to be significant, but less important as a determinant of major choice in France than heterogeneity in preferences for the subject matter. We revisit the issue using our nonparametrics bounds methodology on Canadian and German data. We examine whether the data is consistent with a Roy selection of students into the two sectors based on anticipated labour market outcomes only. We study how the answers depend on visible minority status and residency in Québec and the former German Democratic Republic.

The picture is rather different for women. The labour market advantages, if present, are not so clear-cut, as noted by Zafar (2013) and Hunt, Garant, Herman, and Munroe (2013), and women are severely under-represented in STEM education and even more so in STEM jobs. The evidence is summarized in Beede, Julian, Langdon, McKittrick, Khan, and Doms (2011). Two dominant explanations for the under-representation of women in STEM education and in STEM careers are discrimination, which lowers expected wages for women in STEM, and gender profiling, which keeps young women away from STEM education. The former is compatible with a Roy model of career choice, assuming wage discrimination is anticipated, and can be addressed by policies fighting lower labour market outcomes for women in STEM. The latter involves non pecuniary considerations in major choice, therefore requires generalized Roy modeling and can be addressed by policies aimed at encouraging young women into STEM education. However, differential costs between STEM and non STEM majors are nonexistent in Germany and Canada, so that generalized Roy models based on differential costs are not directly applicable here. In any event, given the divergence in policy implications of the two channels above, it is important to investigate which of the two is the dominant effect.

The under-representation of women in STEM jobs is often cited as a major contributor to the gender wage gap, as in Daymont and Andrisani (1984). More generally, there is a large amount of informal discussion, although, to the best of our knowledge, little formal investigation, of the contribution of the STEM economy to rising wage inequality; see, for instance, Brynjolfsson and McAfee (2012), who attribute rising inequality to skill-based technological change. Our methodology allows us to address this issue by comparing inequality in STEM wages to inequality in non-STEM wages in a counterfactual economy, where sector allocation is random. We can also investigate the effect of self-selection on sectoral and aggregate wage inequality.

Data.

Our empirical analysis relies on Canadian and German National Graduates Surveys. Both contain detailed information on a representative sample of recent university graduates in their respective countries. The German data are collected by the German Centre for Higher Education Research and Science Studies (DZHW) as part of the DZHW Graduate Survey Series. Data and methodology are described in Baillet, Franken, and Weber (2017). In Germany, the wave we consider includes graduates who obtained their highest degree during the academic year 2008-2009, whereas in Canada, the wave we consider includes graduates who obtained their highest degree during the academic year 2009-2010. We also examine data from earlier waves, namely 1997, 2001, 2005 for Germany and 2000, 2005 for Canada. In the case of the earlier Canadian waves, we rely on publicly available data, which has fewer variables than the administrator data we use for the 2009-2010 wave, and which, unlike the latter, only provides interval censored income information.

Graduates were interviewed 1 year and 5 years after graduation in the German survey and 3 years after graduation in the Canadian survey. At that point, extensive information was collected on their educational experience, employment history, including wages and hours worked, along with detailed socio-economic variables. Geographical information is more precise in the German data, with 38 regions, as opposed to 13 in the Canadian data. The German data also contains information on talent, with results at the *Abitur* (high school final exam), whereas the Canadian survey only provides a self-assessed measure of ability.

Both data sets allow us to observe whether employment is permanent or temporary and whether it is related to the specific field of study. In both data sets, fields of study are recorded at a high level of disaggregation, which allows us to discriminate subjects that require mathematics from those that don't. We then merge the fields of study into two categories. We call STEM the mathematics intensive category, which consists mostly of mathematic, physical, economic and computer sciences, as well as engineering and related fields, although other STEM definitions often include life sciences and exclude economics. The remaining majors are merged in the non-STEM-degree category. In Canada and in Germany, the choice of field of study is made prior to enrolment in the program. In

both countries, we only consider graduates from institutions in the country of the survey, who are active on their respective country’s labor market at the time of the interview.

We consider a selection of outcome variables: the ability to secure a permanent employment, the ability to secure employment within the field of study, and annual wage and average hourly wage during the year prior to the time of the interview. Given the high correlation between wage and hourly wage measures, we report only results for wage. Annual wage is non censored in the Canadian data and reported in 1,000 euro bins in the German data.

The potential instruments we consider are the education level of both parents (the surveys report parental education in discrete categories, which we translate into years of education following Parey, Ruhose, Waldinger, and Netz (2017)), the proportion of women among STEM faculty members (which we call rate of feminization of the STEM faculty) in universities in the individual’s region of residence at the time of choice. The Canadian version of this variable is drawn from a separate data source, namely Statistics Canada, University and College Academic Staff System (UCASS). There is a very high level of assortative matching in parent’s education both in Canada and Germany, so we only report results using the mother’s education and the rate of feminization of STEM faculty as stochastically monotone instrumental variables. We also use local labor market conditions at the time of choice as instruments for robustness purposes, although their validity relies on neglecting general equilibrium effects.

We compare results for gender and visible minority status. In the Canadian survey, visible minority status is self-reported. In Germany, we construct this variable from the country of birth, and we assign an individual in the survey the status of visible minority if they were born in a country with a non-white majority population. This unfortunately excludes a large number of graduates of Turkish descent, whom we are unable to track. We also distinguish German graduates from institutions in the former German Democratic Republic and Canadian graduates from institutions in Québec.

Our study focuses on the latest cohort. The raw sample from the German survey consists of 10,494 individuals. From the raw sample, we exclude all respondents who are still in education, have never worked or are currently inactive, unemployed, in part-time employment or self-employed. This leaves 9,202 observations. We keep only graduates who hold a “Bachelor”, “Magister” or “Diplom”, excluding those with “Staatsexamen” and “Lehramt” degrees, which are specific tracks mainly for teachers. This leaves us with 7,729 observations. Finally, we divide the population between those who completed the *Abitur* (high school final exam) in the former Federal Republic and in the former Democratic Republic and exclude those for which we do not have this information or obtained their *Abitur* abroad (107 individuals). Most of our econometric analysis is based on the sample of individual with complete information on gender, degree, migration background, year and place of

Abitur completion, mother’s education, and income or job characteristics, that is between 4,559 and 4,890 observations.

The raw sample from the Canadian survey consists of 28,715 observations who participated in the survey. From the raw sample, we exclude all respondents who have completed trade, vocational, college and CEGEP diploma or certificate at the time of their 2009/2010 graduation and 2013 interview. We keep only those individuals who have “university diploma or certificate below Bachelor level”, “Bachelors degree or first professional degree”, “university diploma/certificate above the Bachelors level but below the Masters level”, “Masters degree” and “Doctorate”. We also exclude all respondents who are still in education, self-employed, working in family business without pay and live in the U.S. as primary residence. We further filter the data set to include respondents who are in the labour force, employed, work full-time and have age below 40. Since the econometric analysis is based on the sample with complete information on gender, minority status, income, degree, related job, permanent job, mothers education and fathers education, the sample size ranges between 4,361 and 10,150 observations.

Descriptive statistics.

Income distributions in Germany and Canada show a clear STEM advantage for both men and women and a clear gender gap. In Figure 3, distributions appear to be stochastically ordered. In both Canada and Germany, based on quartiles only, the distribution of male STEM wages dominates the distribution of female STEM wages, which dominates male non STEM wages, which dominates female non STEM wages. A similar pattern emerges from Table 3, where we see that men with STEM degrees are more likely to hold permanent employment in a field related to their studies, than men with non STEM degrees and women in both categories. More precisely, in Germany, 41% of men with STEM degrees obtain permanent employment one year after graduation, 35% in a field related to their studies and 6% in other fields. For women with STEM degrees, the proportion is only 36%, with 29% in their field of study, and for men and women with non STEM degrees, the proportion falls to 23%, with 16% in their field of study. In Canada, 90% of men with STEM degrees obtain permanent employment one year after graduation, 84% in a field related to their studies and 6% in other fields. For women with STEM degrees, the proportion is 82%, with 73% in their field of study, and for men and women with non STEM degrees, the proportion is 81%, with 68% in their field of study. Since the proportion of men with STEM degrees is larger, the overall proportion of women with a permanent employment after 1 year in Germany is lower (27%) than for men (36%) and the proportion of women with a permanent employment after 3 years in Canada is lower (80%) than for men (86%).

Table 2 shows the degree of under representation of women in STEM degrees in both Germany and Canada, which tallies with the overwhelming evidence from previous studies in different contexts. In Germany, 37% of women’s degrees are in STEM, as opposed to 75% for men. The difference is

somewhat less pronounced for minorities, where 48% of women’s degrees are in STEM, as opposed to 80% for men. In Canada, 8% of women’s degrees are in STEM, as opposed to 35% for men. The difference is, again, less pronounced for minorities, where 15% of women’s degrees are in STEM, as opposed to 45% for men.

The distribution of the proportion of women on the STEM faculty in the graduate’s region at the time choice is given in Figure 5. We examine the variation in sector choice induced by the instruments in Figure 4. There is some indication of a hump-shaped response of STEM choices in mother’s education for men and women. The humped-shaped response to mother’s education may be due to a larger involvement in major choices for parents with a bachelor’s degree and a more laissez-faire approach beyond that. The effect of the proportion of women on the STEM faculty on women’s choices is increasing for low proportions, as we would expect, then levels for larger proportions.

Results.

From the survey samples, we first construct sub-samples based on gender, visible minority status, and the broad region of residence at the time of the interview (former East and West Germany, Québec and the rest of Canada). We test monotonicity of the conditional mean for each binary outcome, and both mean monotonicity and stochastic monotonicity of the non-binary discrete and continuous outcomes with respect to the instruments. The tests are adapted from Hsu, Liu, and Shi (2016), to allow for monotonicity with respect to a vector of instrumental variables.¹ The sensitivity of inference results to the generalized moment selection procedure is usually the major concern with this type of procedure. We choose the recommended values for the user-chosen parameters governing the generalized moment selection in Hsu, Liu, and Shi (2016), namely $B_n = 0.85 \ln n / \ln \ln n$ and $\kappa_n = 0.15 \ln n$. To investigate robustness of the inference results to variations around this choice, we ran the tests in the case of the mother’s education as an instrument in Canada for all pairs of values in $\{B_n/2, B_n, 2B_n\} \times \{\kappa_n/2, \kappa_n, 2\kappa_n\}$. Of the 48 test results in the Canadian portions of Tables 5 and 6, only one is affected, namely the level of the rejection of Roy self-sorting behavior based on expected relatedness of employment for white women in Québec goes from 5% to 10%. Inference on the bounds from Propositions 2 and 3 on the efficiency loss from non maximizing behavior and on the bounds (3.2) and (3.4) on the interquartile range in the randomized economy, is carried out with the STATA package *clrbounds* implementing Chernozhukov, Lee, and Rosen (2013).

Table 5 collects results of the test of the Roy model with imperfect foresight using the mother’s education as an instrument satisfying Assumption 5 (SMIV). The hypotheses that white men and women in the former West German Federal Republic choose their major to maximize expected income or the probability of a permanent employment a year after graduation are both rejected at the 1% level. The hypothesis that white women in the former Democratic Republic choose their major to

¹We thank Yu-Chin Hsu and Chu-An Liu for sharing their code.

maximize the probability of a permanent employment a year after graduation is also rejected at the 10% level. No other rejection of imperfect foresight Roy selection are found for residents of the former Democratic Republic or for minorities.

The hypotheses that white women in the rest of Canada choose their major to maximize expected income or the probability of a permanent employment three years after graduation are both rejected at the 5% level. The hypothesis that white women in Québec choose their major to maximize the probability of securing employment related to their field of study is also rejected at the 5% level. For men, Roy self-sorting is never rejected, which again shows a significantly different behavior for men and women. As in Germany, we find no rejections for visible minority men or women. A notable feature of the results presented in Table 5 is that the hypothesis that white Canadian women's choices are driven by expected income or the probability of securing permanent employment is rejected for the rest of Canada, but not in Québec, whereas the hypothesis that choices are driven by the probability of securing employment related to the field of study is rejected in Québec, but not in the rest of Canada. This is consistent with the interpretation that labor market outcomes are stronger determinants of choices for women in Québec, whereas field preferences are stronger determinants of choice for women in the rest of Canada.²

Table 6 reports results of the test of pure Roy self-sorting behavior based on three outcome variables, namely income, the degree to which employment is related to the field of study, and the vector (permanent, related) with lexicographically ordered components. We no longer include the ability to secure permanent employment, since it is a binary variable, and the tests of pure and imperfect foresight models are identical. As we see in Table 6, the same conclusions hold for the pure Roy selection model, except that the hypothesis that white women in the former Democratic Republic choose their major to maximize expected income a year after graduation is now also rejected at the 1% level, and the hypothesis that white women in Québec choose their major to maximize the probability of securing permanent employment three years after graduation (and in case of ties decide based on relatedness of the employment) is now also rejected at the 10% level. Again, there are no rejections of the pure Roy selection model for minorities anywhere, or for men anywhere in Canada.

Tables 7 and 8 collect similar results to those in Tables 5 and 6, except that Assumption 5 (SMIV) holds for the vector of instruments combining mother's education and the proportion of women on the STEM faculty in the individual's region at the time of choice. Hence, only results for women are presented, since the proportion of women on the STEM faculty is conceived as a valid SMIV

²To investigate this issue further, we tested a Roy model of self-sorting based on a variable equal to 1 when the applicant says they obtained the employment they were hoping for, and zero otherwise. This variable is available in Canada and the test result are identical to those obtained for the test of Roy self-sorting based on the relatedness of employment with field of specialization at university. Interpretation of this result, however, would hinge on a correct interpretation of the variable itself, which we do not have at this point.

for women only. Again, there are no rejections of either the imperfect or the perfect foresight Roy models for minorities.

The hypotheses that white women in the former Federal Republic choose their major to maximize expected income or the probability of a permanent employment a year after graduation are both rejected at the 1% level. The hypothesis that white women in the former Democratic Republic choose their major to maximize their expected income (resp. probability of a permanent employment) a year after graduation is also rejected at the 1% level (resp. 5%). The same results hold for the test of perfect foresight Roy self-sorting. Looking at the Canadian portion of Tables 7 and 8 reveals only slight discrepancies with test results with only the mother’s education as the SMIV. One initially surprising result in the Canadian portion of Table 7 is the fact that Roy self-sorting behavior for white women based on income is no longer rejected when the test is based on the vector of instruments, whereas it was rejected based on mother’s education only (and also based on feminization of STEM only, although we didn’t include the result in the tables). This finding is explained by the very low dependence between the instruments. Indeed, we are testing imperfect foresight Roy by testing mean monotonicity of observed outcome relative to a vector of instruments. Heuristically, rejections are generated by sufficiently many pairs of individuals (i, j) such that both instruments yield higher values for individual i than individual j , yet i has lower realized outcome. If the dependence between instruments is low, there are too few pairs of individuals with instrumental values partially ordered, hence the test has little power to reject the hypothesis.

Our results show a prevalence of rejections of Roy major selection behavior, possibly in favor of non pecuniary considerations, for categories that are generally considered privileged, particularly women, i.e., white women in Canada and white men and women from the former West German Federal Republic. We tend not to reject Roy major selection behavior for all other categories. This is borne out by the differences in responses to a survey question on the importance of labor market considerations on major choice. Table 4 shows that minorities and residents of Québec and the former Democratic Republic of Germany tend to weigh labor market considerations more than their counterparts.

To further investigate rejections of the Roy self-sorting behavior, we compute confidence intervals for the measure of departure from Roy (also called “efficiency loss”) provided in Section 2.5 and 3.2. We report the confidence lower bounds for categories of individuals, for whom the Roy self-sorting behavior was rejected, namely white women in the rest of Canada and white men and women from the former West Germany. For each of these categories, we plot the lower confidence bound as a function of income and the mother’s education to identify regions of values (of income and mother’s education) that are responsible for the rejection of Roy self-sorting. For white men in the former West Germany, we find that rejections are mostly driven by individuals, whose mothers earned postgraduate degrees, whereas for white women in the former West Germany, we find rejections

are driven by lower income women with high school educated mothers and median income women, whose mothers earned a high school degree only or a postgraduate degree.

Finally, we investigate the impact of Roy self-selection on income inequality in the case of individuals for whom the hypothesis of Roy self-sorting is not rejected, i.e., minorities of both genders in Germany and Canada, white women in the former East Germany and Canadian white men. In Figure 7, we report confidence intervals for the partially identified interquartile range of potential non STEM income Y_0 , potential STEM income Y_1 , and aggregate income in an economy where individuals are randomized into sectors, next to the interquartile range for observed distributions of STEM, non STEM and aggregate income distributions. Most results are inconclusive, in the sense that realized interquartile ranges are well within the bounds for potential distributions, except in the case of white men in Québec and white women in the former East Germany, where observed STEM interquartile range coincides with the lower bound on potential interquartile range, which is consistent with Roy's hypothesis that self-sorting can reduce inequality.

5. CONCLUSION

In this paper, we analyzed the Roy model of self-sorting into economic activities on the basis of anticipated outcomes. We stripped the model down to its essential features: we assumed that heterogeneous agents are characterized by a pair of potential outcomes, one for each sector of activity, and that they choose the sector that gives them a strictly higher outcome, leaving choice undetermined in case of ties. We characterized the restrictions this mechanism imposes on the joint distribution of potential outcomes. In other words, we characterized the identified set for the joint distribution of potential outcomes. In the case of binary outcomes, we showed how to represent the identified set in two dimensions on a rescaled 2-simplex. This characterization showed, on the one hand, that the Roy self-sorting mechanism puts non trivial restrictions on joint distributional features of potential outcomes, but, on the other hand, that the identified set is never empty, so that the Roy self-sorting mechanism described is not testable. Testability can be restored using selection shifters that are jointly independent of potential outcomes. However, such shifters are difficult to find in applications, and their usefulness is severely restricted by the Roy self-sorting mechanism, which only lets them affect selection in case potential outcomes are equal. We therefore introduced an extension of the notion of monotone instrumental variable, designed to constrain the joint distribution of potential outcomes, the quantity of interest. We considered (vectors of) variables that affect the vector of potential outcomes monotonically, in the sense of multivariate first order stochastic dominance, and called such (vectors of) variables stochastically monotone instrumental variables (SMIV). We repeated the characterization of the identified set for the joint distribution of potential outcomes under the SMIV assumption, and showed that testing the Roy self-sorting mechanism is equivalent to testing stochastic monotonicity of observed outcomes in the instrument. To alleviate the concern that rejections are due to the assumption that agents are perfectly informed

of their future outcomes, we repeated the exercise with an imperfect foresight version of the model, where agents select sectors based on expectations. Beyond testing the Roy self-sorting mechanism and providing measures of departure from outcome-based decisions, we highlighted another important application of our characterization of the identified set of joint potential outcome distributions, namely the derivation of sharp bounds on the interquantile range of potential outcome distributions to revisit the effects of self selection on inequality in employment outcomes.

We applied our methodology to the analysis of major choices made by graduates of Canadian and German universities based on the national graduate surveys of each of these two countries. We analyzed selection of mathematics-intensive versus other fields of study by graduates within the framework of the Roy model with employment based outcomes that include income 1 and 3 years after graduation, the ability to secure permanent employment with a year of graduation and the extent to which employment secured is related to the field of study. The data supports previous evidence of a labor market advantage of mathematics-intensive fields (STEM), severe under-representation of women in STEM, over-representation of visible minorities in STEM and male labor market advantage in both sectors. We investigated whether selection behavior is consistent with Roy self-sorting on outcomes for categories of graduates by gender, visible minority status and region of residence (former East and former West Germany, Québec and the rest of Canada). To test Roy self-sorting based on employment outcomes, we used parental education level and the proportion of women on the faculty of STEM programs in the region and at the time of choice as stochastically monotone instruments. We found a pattern of rejections of Roy self-sorting based on outcomes for white women in the former Federal Republic of Germany and the rest of Canada, and a lack of rejections for visible minorities and for white women from Québec and white men from all of Canada and the former German Democratic Republic. Confidence intervals for measures of departure from Roy behavior revealed that in the case of white women from the former Federal Republic, for instance, rejection of Roy behavior seems to be driven by lower income women with high school educated mothers and middle income women with postgraduate educated mothers. Among groups, where Roy self-sorting is not rejected, comparisons of interquartile ranges for observed and counterfactual income distributions are inconclusive except in the case of women in the former Democratic Republic, where self-sorting is found not to increase inequality.

The pattern of rejections of Roy self-sorting in major choice points to non labor market related determinants of choice. For instance, our results are consistent with a story involving gender profiling pushing white men in the West of Germany into STEM fields and white women in the West of Germany and in Canada out of STEM fields. They are also consistent with gender profiling being less prevalent in the former communist Germany. However, the results are also consistent with a story involving non pecuniary field preferences driving major choices of more privileged groups in more affluent regions, but not the choices of the more financially constrained. The methodology proposed here should then be construed as a tool for exploratory analysis of the determinants of

major choice prior to a fully structural generalized Roy modeling of preferences, in a context where, unlike the analysis of returns to college, there is no clear cost differential between different choices. Non rejections of Roy self-selection based on labor market outcomes, on the other hand, are a warning that policies directly aimed at increasing the share of women in STEM majors at university may have a (possibly short term) negative effect on the gender gap and wage inequality, and that both upstream (early childhood) and downstream (labor market) interventions are required.

APPENDIX A. PROOFS AND ADDITIONAL RESULTS RELATING TO BINARY OUTCOMES

A.1. Sharp bounds for the binary outcome Roy model.

Statement of Proposition 1. Fix the pair of binary random variables (Y, D) with probability mass function $(q_{00}, q_{01}, q_{10}, q_{11})$, with $q_{ij} := \mathbb{P}(Y = i, D = j)$. The following two statements hold. (1) If the non negative vector $(p_{00}, p_{01}, p_{10}, p_{11}) \in \mathbb{R}^4$ satisfies $p_{00} + p_{01} + p_{10} + p_{11} = 1$, $p_{10} \leq q_{10}$, $p_{01} \leq q_{11}$ and $p_{00} = q_{00} + q_{01}$, then there exists a pair of binary random variables (Y_0, Y_1) such that Assumptions 1 and 2 are satisfied and $\mathbb{P}(Y_0 = 0, Y_1 = 0) = p_{00}$, $\mathbb{P}(Y_0 = 0, Y_1 = 1) = p_{01}$, $\mathbb{P}(Y_0 = 1, Y_1 = 0) = p_{10}$ and $\mathbb{P}(Y_0 = 1, Y_1 = 1) = p_{11}$. (2) Conversely, if the pair of binary random variables (Y_0, Y_1) satisfies Assumptions 1 and 2, then $\mathbb{P}(Y_0 = 1, Y_1 = 0) \leq q_{10}$, $\mathbb{P}(Y_0 = 0, Y_1 = 1) \leq q_{11}$ and $\mathbb{P}(Y_0 = 0, Y_1 = 0) = q_{00} + q_{01}$.

Proof of Proposition 1. Write $p_{ij} := \mathbb{P}(Y_0 = i, Y_1 = j)$ for each $i, j = 0, 1$. The binary outcomes Roy model of Definition 1 can be equivalently defined as a correspondence G between values of observables $(y, d) \in \mathcal{A} := \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ and values of unobservables $(y_0, y_1) \in \mathcal{A}$. The correspondence is defined by its values $G(y, d)$ for each $(y, d) \in \mathcal{A}$, namely $G(1, 1) := \{(1, 1), (0, 1)\}$, $G(1, 0) := \{(1, 1), (1, 0)\}$, $G(0, 1) := \{(0, 0)\}$ and $G(0, 0) := \{(0, 0)\}$. By Theorem 1 of Galichon and Henry (2011), the 14 inequalities $\mathbb{P}((Y_0, Y_1) \in A) \leq \mathbb{P}(G(Y, D) \cap A \neq \emptyset)$ for each $A \subset \mathcal{A}$ provide a collection of sharp bounds for the model defined by the correspondence G . For instance, $A = \{(0, 0)\}$ yields the inequality $p_{00} \leq q_{00} + q_{01}$ and $A = \{(1, 1), (0, 1)\}$ yields the inequality $p_{11} + p_{01} \leq q_{11} + q_{10}$. To prove the result, it suffices to show that all 14 inequalities are implied by $0 \leq p_{10} \leq q_{10}$, $0 \leq p_{01} \leq q_{11}$ and $p_{00} = q_{00} + q_{01}$. The 14 inequalities are listed below. Singleton A 's yield

$$\begin{aligned} p_{11} &\leq q_{11} + q_{10} \\ p_{10} &\leq q_{10} \\ p_{01} &\leq q_{11} \\ p_{00} &\leq q_{01} + q_{00}. \end{aligned} \tag{A.1}$$

Pairs yield

$$\begin{aligned} p_{11} + p_{10} &\leq q_{11} + q_{10} \\ p_{11} + p_{01} &\leq q_{11} + q_{10} \\ p_{11} + p_{00} &\leq 1 \\ p_{10} + p_{01} &\leq q_{11} + q_{10} \\ p_{10} + p_{00} &\leq q_{10} + q_{01} + q_{00} \\ p_{01} + p_{00} &\leq q_{11} + q_{01} + q_{00}. \end{aligned} \tag{A.2}$$

Finally, triplets yield

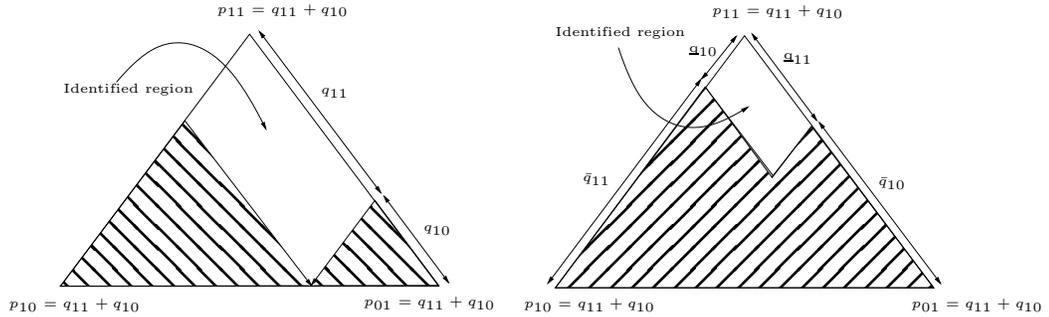
$$\begin{aligned}
 p_{11} + p_{10} + p_{01} &\leq q_{11} + q_{10} \\
 p_{11} + p_{10} + p_{00} &\leq 1 \\
 p_{11} + p_{01} + p_{00} &\leq 1 \\
 p_{10} + p_{01} + p_{00} &\leq 1.
 \end{aligned} \tag{A.3}$$

The first four inequalities in (A.2) are implied by the first inequality in (A.3). The last two are implied by (A.1). All inequalities in (A.2) are therefore redundant. Since $p_{11} = 1 - p_{00} - p_{01} - p_{10}$, all four inequalities in (A.1) are implied by $0 \leq p_{10} \leq q_{10}$, $0 \leq p_{01} \leq q_{11}$ and $p_{00} = q_{00} + q_{01}$. Finally, since $p_{11} + p_{10} + p_{01} = 1 - p_{00}$, the first inequality in (A.3) is implied by $p_{00} = q_{00} + q_{01}$ and the result follows.

Extension to the alternative binary Roy model. (1) First, we show that $\mathbb{P}(Y_0 = 1, Y_1 = 0) \leq q_{10}$, $\mathbb{P}(Y_0 = 0, Y_1 = 1) \leq q_{11}$ and $\mathbb{P}(Y_0 = 0, Y_1 = 0) = q_{00} + q_{01}$ hold if (Y, D, Y_0, Y_1) satisfy the assumptions of Definition 2. Under the specification of Definition 2, $Y_0 = 1$ and $Y_1 = 0$ jointly imply that $Y_0^* > Y_1^*$, which in turn implies $D = 0$ and $Y = 1$, so that the first inequality holds. The second holds by the same reasoning and the roles of Y_0 and Y_1 reversed. Finally, $Y_0 = Y_1 = 0$ implies $Y = 0$, and $Y = 1$ implies that $Y_0 = 1$ or $Y_1 = 1$, so the equality holds as well. (2) Second, the binary outcomes Roy model specification of Definition 1 is nested in the alternative binary Roy model specification of Definition 2. Indeed, the former can be obtained by restricting (Y_0^*, Y_1^*) to be binary. Hence, sharpness of the bounds for the binary outcomes Roy model implies sharpness for the alternative binary Roy model. The result follows.

Representation of the bounds on the 2-simplex. We continue to denote $\mathbb{P}(Y = i, D = j) = q_{ij}$ and $\mathbb{P}(Y_0 = i, Y_1 = j) = p_{ij}$. According to Proposition 1, $p_{00} = q_{01} + q_{00}$. Hence, the remaining three probabilities, namely p_{10} , p_{01} and $p_{11} = q_{11} + q_{10} - p_{10} - p_{01}$ can be represented in barycentric coordinates in the rescaled 2-simplex of Figure 2, where the three vertices correspond to the cases, where $p_{11} = q_{11} + q_{10}$, $p_{10} = q_{11} + q_{10}$ and $p_{01} = q_{11} + q_{10}$ respectively.

Figure 2. Identified set for $(p_{10}, p_{01}, p_{11} = q_{11} + q_{10} - p_{10} - p_{01})$ in barycentric coordinates in the rescaled 2-simplex. p_{00} is identified and equal to $q_{01} + q_{00}$. The left-hand-side figure is without excluded variable Z . The right-hand-side is in the presence of variation in a variable Z satisfying Assumption 5. The conditioning variable z is omitted from the notation in the graph and $\bar{q}_{1j}(z) := \sup_{\tilde{z} \leq z} q_{1j}(\tilde{z})$, $\underline{q}_{1j}(z) := q_{11}(z) + q_{10}(z) - \bar{q}_{1,1-j}(z)$, $j = 0, 1$.



A.2. Covariate restrictions.

Statement of Theorem 1(1). Fix the joint distribution (Y, D, Z) and denote the conditional probability mass function $(q_{00}(z), q_{01}(z), q_{10}(z), q_{11}(z))$, where $q_{ij}(z) = \mathbb{P}(Y = i, D = j | Z = z)$. The following two statements hold. (1) If for each $z \in \text{Supp}(Z)$, the non-negative vector $(p_{00}(z), p_{01}(z), p_{10}(z), p_{11}(z)) \in \mathbb{R}^4$ satisfies $p_{00}(z) + p_{01}(z) + p_{10}(z) + p_{11}(z) = 1$, $\sup_{\tilde{z} \leq z} q_{10}(\tilde{z}) \leq p_{01}(z) + p_{11}(z)$, $\sup_{\tilde{z} \leq z} q_{11}(\tilde{z}) \leq p_{10}(z) + p_{11}(z)$ and $p_{00}(z) = q_{01}(z) + q_{00}(z)$, then there exists a pair of binary random variables (Y_0, Y_1) such that Assumptions 1, 2 and 5 are satisfied and $\mathbb{P}(Y_0 = 0, Y_1 = 0 | Z = z) = p_{00}(z)$, $\mathbb{P}(Y_0 = 0, Y_1 = 1 | Z = z) = p_{01}(z)$ and $\mathbb{P}(Y_0 = 1, Y_1 = 0 | Z = z) = p_{10}(z)$, for each $z \in \text{Supp}(Z)$. (2) Conversely, if the random vector (Y, D, Z) satisfies Assumptions 1, 2 and 5 for some the pair of binary random variables (Y_0, Y_1) , then (2.6) holds.

Proof of Theorem 1(1). From the proof of Proposition 1, the identified set under Assumptions 1 and 2 is characterized by $q_{10}(z) \leq p_{10}(z) + p_{11}(z) \leq q_{10}(z) + q_{11}(z)$, $q_{11}(z) \leq p_{01}(z) + p_{11}(z) \leq q_{10}(z) + q_{11}(z)$ and $p_{00}(z) = q_{01}(z) + q_{00}(z)$ for all $z \in \text{Supp}(Z)$. Assumption 5 is equivalent to $\mathbb{P}((Y_0, Y_1) \in U | Z = z_1) \leq \mathbb{P}((Y_0, Y_1) \in U | Z = z_2)$ for all $z_1 \leq z_2$ and all upper set U in $\{0, 1\}^2$. The upper sets are $\{(1, 1)\}$, $\{(1, 1), (1, 0)\}$, $\{(1, 1), (0, 1)\}$ and $\{(1, 1), (1, 0), (0, 1)\}$. Hence, Assumption 5 is equivalent to $\sup_{\tilde{z} \leq z} p_{11}(\tilde{z}) \leq p_{11}(z) \leq \inf_{\tilde{z} \geq z} p_{11}(\tilde{z})$, $\sup_{\tilde{z} \leq z} [p_{11}(\tilde{z}) + p_{10}(\tilde{z})] \leq p_{11}(z) + p_{10}(z) \leq \inf_{\tilde{z} \geq z} [p_{11}(\tilde{z}) + p_{10}(\tilde{z})]$, $\sup_{\tilde{z} \leq z} [p_{11}(\tilde{z}) + p_{01}(\tilde{z})] \leq p_{11}(z) + p_{01}(z) \leq \inf_{\tilde{z} \geq z} [p_{11}(\tilde{z}) + p_{01}(\tilde{z})]$, and $\sup_{\tilde{z} \leq z} [1 - p_{00}(\tilde{z})] \leq 1 - p_{00}(z) \leq \inf_{\tilde{z} \geq z} [1 - p_{00}(\tilde{z})]$ for all $z \in \text{Supp}(Z)$. Combining the two sets of inequalities yields the result.

Sector specific exclusions. We denote by X_d the vector of observable variables (if any) that enter in the equation determining potential outcome Y_d , but not Y_{1-d} . Since there is some ambiguity in notation, it is worth stressing the fact that both vectors X_0 and X_1 are observed, irrespective of the chosen sector, unlike Y_0 , which is only observed if $D = 0$ and Y_1 , when $D = 1$.

Assumption 6 (Sector specific exclusions). *The random vectors X_0 and X_1 denote vectors of observed variables (when they exist) such that $Y_d \perp\!\!\!\perp X_{1-d} | X_d$, for $d = 0$ and 1.*

The exclusions of Assumption 6 are conditional on a set of additional observed covariates, as noted before. Excluded variables X_d are variables that change the price of skills relevant for one sector without affecting the price of skills in the other, as discussed in Heckman and Honoré (1990). Typical examples would include sector specific shifters of labor market conditions, as in Heckman and Sedlacek (1985, 1990). In the case of college major choice, considered in Section 4, in a narrow partial equilibrium sense, exogenous and unanticipated variation (at the time of college major choice) in the gross number of STEM jobs could be thought to affect only conditions for success in securing employment with a STEM degree, without affecting success in securing employment with a non STEM degree.

The classical way to derive bounds under an exclusion restriction is to observe that $\mathbb{E}(Y_d | X_d, X_{1-d}) = \mathbb{E}(Y_d | X_d)$ under Assumption 6, so that the bounds (2.3) hold for all values of X_{1-d} . We contribute to the literature here, in showing sharpness of these bounds for the binary (and alternative binary) Roy model. Conditioning on all non excluded variables remains implicit throughout.

Proposition 5 (Marginal bounds with sector specific covariates). *For any (x_0, x_1) in the support of (X_0, X_1) , the identified set for the parameter vector $(\mathbb{E}(Y_0|X_0 = x_0), \mathbb{E}(Y_1|X_1 = x_1))$ in the binary (and alternative binary) Roy model is characterized by:*

$$\begin{aligned} \mathbb{P}(Y = 1, D = 0|X_0 = x_0, X_1 = \tilde{x}_1) &\leq \mathbb{E}(Y_0|X_0 = x_0) \leq \mathbb{P}(Y = 1|X_0 = x_0, X_1 = \tilde{x}_1), \\ \mathbb{P}(Y = 1, D = 1|X_0 = \tilde{x}_0, X_1 = x_1) &\leq \mathbb{E}(Y_1|X_1 = x_1) \leq \mathbb{P}(Y = 1|X_0 = \tilde{x}_0, X_1 = x_1), \end{aligned}$$

for almost all $\tilde{x}_1 \in \text{Supp}(X_1|X_0 = x_0)$, and $\tilde{x}_0 \in \text{Supp}(X_0|X_1 = x_1)$.

The bounds define the identified set for the vector $(\mathbb{E}(Y_0|X_0 = x_0), \mathbb{E}(Y_1|X_1 = x_1))$, namely, any value of that vector satisfying the bounds can be achieved as a solution of the model for some distribution of the observable variables (Y, D) conditional on $(X_0 = x_0, X_1 = x_1)$. In other words, no value for the pair $(\mathbb{E}(Y_0|X_0 = x_0), \mathbb{E}(Y_1|X_1 = x_1))$ that satisfies both equations can be rejected solely on the basis of the model specification. The bounds are well-known, but the joint sharpness result is new. As before, the bounds of Proposition 5 are *intersection bounds*, so that inference can be carried out with the method proposed in Chernozhukov, Lee, and Rosen (2013).

A salient consequence of Proposition 5 is the fact that the binary outcomes Roy model can be rejected when the bounds cross, i.e., when there is a value x_1 in the support of X_1 and two values x_0^1 and x_0^2 in the support of X_0 conditional on $X_1 = x_1$, such that $\mathbb{P}(Y = 1, D = 1|X_0 = x_0^1, X_1 = x_1) > \mathbb{P}(Y = 1|X_0 = x_0^2, X_1 = x_1)$ or a value x_0 in the support of X_0 and two values x_1^1 and x_1^2 in the support of X_1 conditional on $X_0 = x_0$, such that $\mathbb{P}(Y = 1, D = 0|X_0 = x_0, X_1 = x_1^1) > \mathbb{P}(Y = 1|X_0 = x_0, X_1 = x_1^2)$. Identification of the pair $(\mathbb{E}(Y_0|X_0 = x_0), \mathbb{E}(Y_1|X_1 = x_1))$ can be achieved as a simple implication of the previous result if there is $\tilde{x}_1 \in \text{Supp}(X_1|X_0 = x_0)$ such that $\mathbb{P}(Y = 1, D = 1|X_0 = x_0, X_1 = \tilde{x}_1) = 0$ and $\tilde{x}_0 \in \text{Supp}(X_0|X_1 = x_1)$ such that $\mathbb{P}(Y = 1, D = 0|X_0 = \tilde{x}_0, X_1 = x_1) = 0$, in which case lower and upper bounds coincide in Proposition 5. This identification result is akin to the *identification at infinity* of Heckman (1990).

Proof of Proposition 5. Validity of the bounds was shown above. For sharpness, fix (x_0, x_1) in the Support of (X_0, X_1) . For a given random vector (Y, D) of binary random variables, denote by $q_{ij}(\tilde{x}_0, \tilde{x}_1)$ the conditional probability $\mathbb{P}(Y = i, D = j|X_0 = \tilde{x}_0, X_1 = \tilde{x}_1)$ for any $(\tilde{x}_0, \tilde{x}_1)$ in the Support of (X_0, X_1) . Consider any pair $(a(x_0), b(x_1))$ satisfying

$$q_{10}(x_0, \tilde{x}_1) \leq a(x_0) \leq q_{11}(x_0, \tilde{x}_1) + q_{10}(x_0, \tilde{x}_1) \tag{A.4}$$

for almost all $\tilde{x}_1 \in \text{Supp}(X_1|X_0 = x_0)$, and

$$q_{11}(\tilde{x}_0, x_1) \leq b(x_1) \leq q_{11}(\tilde{x}_0, x_1) + q_{10}(\tilde{x}_0, x_1) \tag{A.5}$$

for almost all $\tilde{x}_0 \in \text{Supp}(X_0|X_1 = x_1)$. We exhibit a pair of binary random variables (Y_0, Y_1) with joint distribution $p_{ij} := \mathbb{P}(Y_0 = i, Y_1 = j|X_0 = x_0, X_1 = x_1)$, such that Assumptions 1, 2 and 6 are satisfied, and such that

$$p_{11}(x_0, x_1) + p_{10}(x_0, x_1) = a(x_0) \text{ and } p_{11}(x_0, x_1) + p_{01}(x_0, x_1) = b(x_1). \tag{A.6}$$

Here is our proposed distribution.

$$\begin{aligned}
p_{00}(x_0, x_1) &= q_{00}(x_0, x_1) + q_{01}(x_0, x_1), \\
p_{11}(x_0, x_1) &= b(x_1) + a(x_0) - q_{10}(x_0, x_1) - q_{11}(x_0, x_1), \\
p_{10}(x_0, x_1) &= q_{10}(x_0, x_1) + q_{11}(x_0, x_1) - b(x_1), \\
p_{01}(x_0, x_1) &= q_{10}(x_0, x_1) + q_{11}(x_0, x_1) - a(x_0).
\end{aligned}$$

Note that (A.6) is verified by construction. We also verify that $p_{00}(x_0, x_1) + p_{01}(x_0, x_1) + p_{10}(x_0, x_1) + p_{11}(x_0, x_1) = 1$ and that p_{00} , p_{10} , and p_{01} are nonnegative. From (A.4) and (A.5), $q_{10}(x_1, x_0) + q_{11}(x_1, x_0) \leq a(x_0) + b(x_1)$, which implies that $p_{11}(x_0, x_1)$ is also nonnegative. Assumption 6 is implied by (A.6) irrespective of the construction of (Y_0, Y_1) . We now construct a pair (Y_0, Y_1) with conditional distribution $p_{ij}(x_0, x_1)$ such that Assumptions 1 and 2 are both satisfied. First construct a random variable U with uniform distribution on $[0, 1]$ in the following way. Set $U \in [0, q_{00}(x_0, x_1) + q_{01}(x_0, x_1)]$ if and only if $Y = 0$. Set $U \in (q_{00}(x_0, x_1) + q_{01}(x_0, x_1), q_{00}(x_0, x_1) + q_{01}(x_0, x_1) + q_{10}(x_0, x_1)]$ if and only if $(Y, D) = (1, 0)$. Finally, set $U \in (q_{00}(x_0, x_1) + q_{01}(x_0, x_1) + q_{10}(x_0, x_1), 1]$ if and only if $(Y, D) = (1, 1)$. Now set $(Y_0, Y_1) = (0, 0)$ if and only if $U \leq q_{00}(x_0, x_1) + q_{01}(x_0, x_1)$, $(Y_0, Y_1) = (1, 0)$ if and only if $U \in (q_{00}(x_0, x_1) + q_{01}(x_0, x_1), q_{00}(x_0, x_1) + q_{01}(x_0, x_1) + p_{10}(x_0, x_1)]$, $(Y_0, Y_1) = (1, 1)$ if and only if $U \in (q_{00}(x_0, x_1) + q_{01}(x_0, x_1) + p_{10}(x_0, x_1), 1 - p_{01}(x_0, x_1)]$, and $(Y_0, Y_1) = (0, 1)$ if and only if $U \in (1 - p_{01}(x_0, x_1), 1]$. By construction, (Y_0, Y_1) has probability mass distribution $p_{ij}(x_0, x_1)$ and satisfies Assumptions 1 and 2. This completes the proof.

Statement of Theorem 2(1). Fix the joint distribution (Y, D, Z) and denote the conditional probability mass function $(q_{00}(z), q_{01}(z), q_{10}(z), q_{11}(z))$, where $q_{ij}(z) = \mathbb{P}(Y = i, D = j | Z = z)$. The following two statements hold. (1) If for each $z \in \text{Supp}(Z)$, the non-negative vector $(p_{00}(z), p_{01}(z), p_{10}(z), p_{11}(z)) \in \mathbb{R}^4$ satisfies $p_{00}(z) + p_{01}(z) + p_{10}(z) + p_{11}(z) = 1$, $\sup_{\tilde{z} \leq z} q_{10}(\tilde{z}) \leq p_{10}(z) + p_{11}(z) \leq \inf_{\tilde{z} \geq z} [q_{11}(\tilde{z}) + q_{10}(\tilde{z})]$, $\sup_{\tilde{z} \leq z} q_{11}(\tilde{z}) \leq p_{01}(z) + p_{11}(z) \leq \inf_{\tilde{z} \geq z} [q_{11}(\tilde{z}) + q_{10}(\tilde{z})]$, $p_{00}(z) \leq \inf_{\tilde{z} \leq z} [q_{01}(\tilde{z}) + q_{00}(\tilde{z})]$, $p_{10}(z) \leq q_{01}(z) + q_{10}(z)$, and $p_{01}(z) \leq q_{00}(z) + q_{11}(z)$, then there exists a pair of binary random variables (Y_0, Y_1) such that Assumptions 1, 3 and 5 are satisfied and $\mathbb{P}(Y_0 = 0, Y_1 = 0 | Z = z) = p_{00}(z)$, $\mathbb{P}(Y_0 = 0, Y_1 = 1 | Z = z) = p_{01}(z)$ and $\mathbb{P}(Y_0 = 1, Y_1 = 0 | Z = z) = p_{10}(z)$, for each $z \in \text{Supp}(Z)$. (2) Conversely, if the random vector (Y, D, Z) satisfies Assumptions 1, 3 and 5 for some the pair of binary random variables (Y_0, Y_1) , then (2.7) and (2.8) hold.

Statement of Proposition 2(1). Fix the joint distribution (Y, D, Z) and denote the conditional probability mass function $(q_{00}(z), q_{01}(z), q_{10}(z), q_{11}(z))$, where $q_{ij}(z) = \mathbb{P}(Y = i, D = j | Z = z)$. The following two statements hold. (1) If for each $z \in \text{Supp}(Z)$, the non-negative vector $(p_{00}(z), p_{01}(z), p_{10}(z), p_{11}(z)) \in \mathbb{R}^4$ satisfies $p_{00}(z) + p_{01}(z) + p_{10}(z) + p_{11}(z) = 1$, $\sup_{\tilde{z} \leq z} q_{10}(\tilde{z}) \leq p_{10}(z) + p_{11}(z) \leq 1 - \sup_{\tilde{z} \geq z} q_{00}(\tilde{z})$, $\sup_{\tilde{z} \leq z} q_{11}(\tilde{z}) \leq p_{01}(z) + p_{11}(z) \leq 1 - \sup_{\tilde{z} \geq z} q_{01}(\tilde{z})$, $p_{11}(z) \leq \inf_{\tilde{z} \geq z} [q_{10}(\tilde{z}) + q_{11}(\tilde{z})]$, $p_{00}(z) \leq \inf_{\tilde{z} \leq z} [q_{01}(\tilde{z}) + q_{00}(\tilde{z})]$, $p_{10}(z) \leq q_{01}(z) + q_{10}(z)$, and $p_{01}(z) \leq q_{00}(z) + q_{11}(z)$, then there exists a pair of binary random variables (Y_0, Y_1) such that Assumptions 1 and 5 are satisfied and $\mathbb{P}(Y_0 = 0, Y_1 = 0 | Z = z) = p_{00}(z)$, $\mathbb{P}(Y_0 = 0, Y_1 = 1 | Z = z) = p_{01}(z)$ and $\mathbb{P}(Y_0 = 1, Y_1 = 0 | Z = z) = p_{10}(z)$, for each $z \in \text{Supp}(Z)$. (2) Conversely, if the random vector (Y, D, Z) satisfies Assumptions 1 and 5 for some the pair of binary random variables (Y_0, Y_1) , then (2.9) and (2.10) hold.

Proof of Proposition 2(1). For each z in the support of Z , the binary outcomes model under Assumption 1 can be equivalently defined as a correspondence G between values of observables $(y, d) \in \mathcal{A} := \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ and values of unobservables $(y_0, y_1) \in \mathcal{A}$. The correspondence is defined by its values $G(y, d)$ for each $(y, d) \in \mathcal{A}$, namely $G(1, 1) := \{(1, 1), (0, 1)\}$, $G(1, 0) := \{(1, 1), (1, 0)\}$, $G(0, 1) := \{(1, 0), (0, 0)\}$ and $G(0, 0) := \{(0, 1), (0, 0)\}$. By Theorem 1 of Galichon and Henry (2011), the 14 inequalities $\mathbb{P}((Y_0, Y_1) \in A|Z = z) \leq \mathbb{P}(G(Y, D) \cap A \neq \emptyset|Z = z)$ for each $A \subset \mathcal{A}$ provide a collection of sharp bounds for the model defined by the correspondence G . The 14 inequalities are listed below. Singleton A 's yield

$$\begin{aligned} p_{11}(z) &\leq q_{11}(z) + q_{10}(z) \\ p_{10}(z) &\leq q_{10}(z) + q_{01}(z) \\ p_{01}(z) &\leq q_{11}(z) + q_{00}(z) \\ p_{00}(z) &\leq q_{01}(z) + q_{00}(z). \end{aligned} \tag{A.7}$$

Pairs yield

$$\begin{aligned} p_{11}(z) + p_{10}(z) &\leq q_{11}(z) + q_{10}(z) + q_{01}(z) \\ p_{11}(z) + p_{01}(z) &\leq q_{11}(z) + q_{10}(z) + q_{00}(z) \\ p_{11}(z) + p_{00}(z) &\leq 1 \\ p_{10}(z) + p_{01}(z) &\leq 1 \\ p_{10}(z) + p_{00}(z) &\leq q_{10}(z) + q_{01}(z) + q_{00}(z) \\ p_{01}(z) + p_{00}(z) &\leq q_{11}(z) + q_{01}(z) + q_{00}(z). \end{aligned} \tag{A.8}$$

Finally, triplets yield only trivial inequalities of the form $p_{11}(z) + p_{10}(z) + p_{01}(z) \leq 1$.

All non trivial inequalities in (A.8) are equivalent to $q_{10}(z) \leq p_{10}(z) + p_{11}(z) \leq 1 - q_{00}(z)$ and $q_{11}(z) \leq p_{01}(z) + p_{11}(z) \leq 1 - q_{01}(z)$. Combining with Assumption 5 as in the proof of Theorem 1(1) and removing redundant inequalities yields the result.

Proof of Theorem 2(1). Assumption 3 is equivalent to $Y = Y_d \Rightarrow \mathbb{E}[Y|Z] = \mathbb{E}[Y|Z] \geq \mathbb{E}[Y_{1-d}|Z]$ for $d = 0, 1$. The latter statement is true for some σ -algebra that contains $\sigma(Z)$ if and only if $Y = Y_d \Rightarrow \mathbb{E}[Y|Z] = \mathbb{E}[Y_d|Z] \geq \mathbb{E}[Y_{1-d}|Z]$ for $d = 0, 1$. The latter is equivalent to $\max\{p_{01}(z) + p_{11}(z), p_{10}(z) + p_{11}(z)\} \leq q_{11}(z) + q_{10}(z)$ for all $z \in \text{Supp}(Z)$. Combining with (A.7) and (A.8) and removing redundant inequalities, yields $q_{10}(z) \leq p_{01}(z) + p_{11}(z) \leq q_{11}(z) + q_{10}(z)$, $q_{11}(z) \leq p_{10}(z) + p_{11}(z) \leq q_{11}(z) + q_{10}(z)$, $p_{00}(z) \leq q_{01}(z) + q_{00}(z)$, $p_{10}(z) \leq q_{01}(z) + q_{10}(z)$, and $p_{01}(z) \leq q_{00}(z) + q_{11}(z)$ for all $z \in \text{Supp}(Z)$. Combining with Assumption 5 as in the proof of Theorem 1(1) yields the result.

Combined sector-specific and SMIV instrument. Suppose Z satisfies Assumption 5 (SMIV) and $Y_0 \perp\!\!\!\perp Z$, so that Z is both a stochastically monotone instrument and a sector specific variable in the sense that it does not directly affect potential outcomes in the non STEM sector. Then, the joint distribution of potential outcomes in the binary outcomes Roy model (Assumptions 1 and 2) satisfies

$$\begin{aligned} \sup_z \mathbb{P}(Y = 1, D = 0|Z = z) &\leq \mathbb{P}(Y_0 = 1) \leq \inf_z \mathbb{P}(Y = 1|Z = z) \\ \sup_{z \leq \tilde{z}} \mathbb{P}(Y = 1, D = 1|Z = \tilde{z}) &\leq \mathbb{P}(Y_1 = 1|Z = z) \\ \mathbb{P}(Y_0 = Y_1 = 0|Z = z) &= \mathbb{P}(Y = 0|Z = z). \end{aligned}$$

Testable implications are stochastic monotonicity of Y relative to Z and

$$\sup_z \mathbb{P}(Y = 1, D = 0|Z = z) \leq \inf_z \mathbb{P}(Y = 1|Z = z).$$

APPENDIX B. PROOFS AND ADDITIONAL RESULTS RELATING TO MIXED DISCRETE-CONTINUOUS OUTCOMES

B.1. Functionally sharp bounds for the Roy model. We first illustrate functional sharpness by showing improvements over Peterson bounds. Combining Peterson bounds (3.1) and assuming $y_{12} > y_{01}$ and $y_{02} > y_{11}$ yields the following upper bound.

$$\begin{aligned}
\mathbb{P}(y_{01} < Y_0 \leq y_{02}, y_{11} < Y_1 \leq y_{12}) &= \mathbb{P}(Y_0 \leq y_{02}, Y_1 \leq y_{12}) - \mathbb{P}(Y_0 \leq y_{02}, Y_1 \leq y_{11}) \\
&\quad - \mathbb{P}(Y_0 \leq y_{01}, Y_1 \leq y_{12}) + \mathbb{P}(Y_0 \leq y_{01}, Y_1 \leq y_{11}) \\
&\leq \mathbb{P}(Y \leq y_{02}, D = 0) + \mathbb{P}(Y \leq y_{12}, D = 1) \\
&\quad - \mathbb{P}(Y \leq \min(y_{02}, y_{11})) \\
&\quad - \mathbb{P}(Y \leq \min(y_{01}, y_{12})) \\
&\quad + \mathbb{P}(Y \leq y_{01}, D = 0) + \mathbb{P}(Y \leq y_{11}, D = 1) \\
&= \mathbb{P}(y_{11} < Y \leq y_{02}, D = 0) + \mathbb{P}(y_{01} < Y \leq y_{12}, D = 1).
\end{aligned}$$

The latter bounds are not sharp. Indeed:

$$\begin{aligned}
\mathbb{P}(y_{01} < Y_0 \leq y_{02}, y_{11} < Y_1 \leq y_{12}) &= \mathbb{P}(y_{01} < Y_0 \leq y_{02}, y_{11} < Y_1 \leq y_{12}, Y_1 \leq Y_0) \\
&\quad + \mathbb{P}(y_{01} < Y_0 \leq y_{02}, y_{11} < Y_1 \leq y_{12}, Y_1 > Y_0) \\
&\leq \mathbb{P}(\max(y_{01}, y_{11}) < Y \leq y_{02}, D = 0) \\
&\quad + \mathbb{P}(\max(y_{01}, y_{11}) < Y \leq y_{12}, D = 1),
\end{aligned}$$

obtained from Theorem 3 (or directly), are sharper unless $y_{01} = y_{11}$.

Proof of Theorem 3(1,2). The Roy model defined by Assumptions 1 and 2 can be equivalently recast as a correspondence $G : \mathbb{R} \times \{0, 1\} \rightrightarrows \mathbb{R}^2$ defined as follows, with the order convention (Y_0, Y_1) for the pair of unobserved variables. For all $y \in \mathbb{R}$,

$$\begin{aligned}
G(y, 0) &= \{y\} \times [\underline{b}, y] \\
G(y, 1) &= [\underline{b}] \times \{y\}.
\end{aligned} \tag{B.1}$$

Indeed, if $D = 0$, by Assumption 1, $Y_0 = Y$. By Assumption 2, $Y_1 \leq Y$. Hence the set of values compatible with the Roy model specification is $(Y_0, Y_1) \in \{Y\} \times [\underline{b}, Y]$, as in the definition of G . Similarly, if $D = 1$, by Assumption 1, $Y_1 = Y$. By Assumption 2, $Y_0 \leq Y$. Hence the set of values compatible with the Roy model specification is $(Y_0, Y_1) \in [\underline{b}, Y] \times \{Y\}$.

The collection (μ, G, ν) , where μ is the joint distribution of the vector (Y, D) of observable variables and ν is the joint distribution of the vector (Y_1, Y_0) of unobservable variables, forms a structure in the terminology of Koopmans and Reiersol (1950) extended by Jovanovic (1989). The correspondence G is non-empty valued and measurable, in the sense that for any open set $\mathcal{O} \subseteq \mathbb{R}^2$, $G^{-1}(\mathcal{O}) := \{(y, d) \in \mathbb{R} \times \{0, 1\} \mid G(y, d) \cap \mathcal{O} \neq \emptyset\}$ is a Borel subset of $\mathbb{R} \times \{0, 1\}$. Hence Theorem 1 of Galichon and Henry (2011) applies and the collection

of inequalities

$$\mu(A) \leq \nu[G^{-1}(A)], \text{ for all Borel } A \subseteq \mathbb{R}^2$$

define sharp bounds for the joint distribution ν of the unobservable variables (Y_1, Y_0) .

For any Borel $A \subseteq \mathbb{R}^2$,

$$\begin{aligned} G^{-1}(A) &= \{(y, d) \in \mathbb{R} \times \{0, 1\} \mid G(y, d) \cap A \neq \emptyset\} \\ &= \{(y, 0) \mid y \in U_{A,0}\} \cup \{(y, 1) \mid y \in U_{A,1}\}. \end{aligned}$$

Hence, $\mu(A) \leq \nu[G^{-1}(A)]$ is equivalent to the second inequality in the display of Proposition 1. The first inequality in that same display is obtained by complementation as follows.

$$\begin{aligned} \mu(A^c) \leq \nu[G^{-1}(A^c)] &\Rightarrow \mu(A) \geq \nu[\{(y, d) \in \mathbb{R} \times \{0, 1\} \mid G(y, d) \subseteq A\}] \\ &= \nu[\{(y, 0) \mid y \in L_{A,0}\} \cup \{(y, 1) \mid y \in L_{A,1}\}] \end{aligned}$$

as required.

Proof of Theorem 3(3). Let A be an upper set in \mathbb{R}^2 . By Assumption 5, $\mathbb{P}((Y_0, Y_1) \in A \mid Z = z) \leq \inf_{\tilde{z} \geq z} \mathbb{P}((Y_0, Y_1) \in A \mid Z = \tilde{z})$. As shown in the proof of Theorem 3(1,2), $\mathbb{P}((Y_0, Y_1) \in A \mid Z = \tilde{z}) \leq \mathbb{P}(Y \in U_{A,0}, D = 0 \mid Z = \tilde{z}) + \mathbb{P}(Y \in U_{A,1}, D = 1 \mid Z = \tilde{z})$. Since A is an upper set, $Y \in U_{A,d}$ implies $(Y, Y) \in A$ and the upper bound follows. Similarly, we have $\mathbb{P}((Y_0, Y_1) \in A \mid Z = z) \geq \mathbb{P}(Y \in L_{A,0}, D = 0 \mid Z = \tilde{z}) + \mathbb{P}(Y \in L_{A,1}, D = 1 \mid Z = \tilde{z})$ for any $\tilde{z} \leq z$. By definition of $L_{A,0}$, $Y \in L_{A,0}$ implies $Y \times [\underline{b}, Y] \subseteq A$. Since A is an upper set, this in turn implies that $[Y, \infty) \times \mathbb{R} \subseteq A$, hence that $Y \geq \underline{y}_0^A$. Reasoning identically for $L_{A,1}$ yields the lower bound and the result follows.

Statement of Corollary 1.

- (1) Let (Y_0, Y_1) be an arbitrary pair of random variables. Let Y and D satisfy Assumptions 1 and 2. Then the distribution functions F_1 and F_0 of Y_1 and Y_0 respectively, satisfy

$$F_d(y_2) - F_d(y_1) \geq \mathbb{P}(y_1 < Y \leq y_2, D = d) + \mathbb{P}(Y \leq y_2, D = 1 - d)1\{y_1 = \underline{b}\} \quad (\text{B.2})$$

for $d = 0, 1$, and for all y_1 and y_2 in $\mathbb{R} \cup \{\pm\infty\}$, such that $y_1 < y_2$.

- (2) Let Y be an arbitrary random variable and D be a binary random variable. Let F_1 and F_0 be right-continuous functions satisfying (B.2). Then there exists a pair (Y_1, Y_0) with cdfs F_1 and F_0 respectively, such that Assumptions 1 and 2 are satisfied.

Proof of Corollary 1.

- (1) Validity of the bounds: As shown in the main text, Proposition 1 yields bounds (3.2)-(3.3) and (3.3) is redundant. Hence the result.

- (2) Sharpness of the bounds: Let Y and D be given. Let F_1 and F_0 be right-continuous functions satisfying (B.2). We shall construct a pair (Y_0, Y_1) with cdfs F_0 and F_1 respectively, such that Assumptions 1 and 2 are satisfied. First, we verify that F_d is a valid cdf for $d = 0, 1$. (B.2) implies that F_d is non decreasing and

non negative. Fixing y_2 and letting $y_1 \rightarrow -\infty$ also yields $F_d(-\infty) = 0$. Finally, letting $y_2 \rightarrow +\infty$ yields $F_d(+\infty) = 1$.

Define \underline{F}_d with $y \mapsto \underline{F}_d(y) = \mathbb{P}(Y \leq y, D = d)$ for each y . Let F^{-1} be the generalized inverse, defined as $F^{-1}(u) = \inf\{y : F(y) \geq u\}$. Let U be a uniform random variable on $[0, 1]$ such that $U < \mathbb{P}(D = 1) \Leftrightarrow D = 1$. Define Y_0 and Y_1 in the following way. When $U < \mathbb{P}(D = 1)$, let $Y_1 = \underline{F}_1^{-1}(U)$ and $Y_0 = (F_0 - \underline{F}_0)^{-1}(U)$. The latter is well defined, since U remains in the range of $F_0 - \underline{F}_0$. Indeed, (B.2) implies

$$F_d(y) \geq \underline{F}_d(y) + \underline{F}_{1-d}(y), \text{ for each } y \in \bar{\mathbb{R}}. \quad (\text{B.3})$$

Hence, $F_d(y) - \underline{F}_d(y) \geq \mathbb{P}(Y \leq y, D = 1 - d)$, hence, in particular, $(F_0 - \underline{F}_0)(+\infty) \geq \mathbb{P}(D = 1)$. For $U > \mathbb{P}(D = 1)$, let $Y_0 = \underline{F}_0^{-1}(U - \mathbb{P}(D = 1))$ and $Y_1 = (F_1 - \underline{F}_1)^{-1}(U - \mathbb{P}(D = 1))$. The latter is well defined because, as before, $(F_1 - \underline{F}_1)(+\infty) \geq \mathbb{P}(D = 0)$.

We first verify Assumption 2. Note first that Assumption 2 is equivalent to $D = d \Rightarrow Y_d \geq Y_{1-d}$ for $d = 0, 1$. Hence, we need to show that $U < \mathbb{P}(D = 1) \Rightarrow Y_1 \geq Y_0$ and $U > \mathbb{P}(D = 1) \Rightarrow Y_1 \leq Y_0$. By symmetry, we only show the first implication. Suppose $U < \mathbb{P}(D = 1)$. If U is a continuity value of \underline{F}_1 , then $U = \underline{F}_1(Y_1)$. By (B.3), $\underline{F}_1 \leq F_0 - \underline{F}_0$. Hence, $U = \underline{F}_1(Y_1) \leq (F_0 - \underline{F}_0)(Y_1)$. So if we can show right-continuity and monotonicity of $F_d - \underline{F}_d$, then $Y_0 = (F_0 - \underline{F}_0)^{-1}(U) \leq Y_1$ as required. Now, monotonicity of $F_d - \underline{F}_d$ follows immediately from (B.2) and right continuity of $F_d - \underline{F}_d$ from that of F_d and \underline{F}_d . If the distribution of Y_1 has an atom at $\underline{F}_1^{-1}(U)$, then, by right-continuity of \underline{F}_1 , $U \leq \underline{F}_1(Y_1) \leq (F_0 - \underline{F}_0)(Y_1)$, so that, by right continuity and monotonicity of $F_0 - \underline{F}_0$, we have $Y_0 = (F_0 - \underline{F}_0)^{-1}(U) \leq Y_1$ as required.

We now verify Assumption 1. We need to show that for each $d = 1, 0$, $\mathbb{P}(Y_d \leq y, D = d) = \underline{F}_d(y)$. By symmetry, we only deal with Y_1 . By monotonicity and right continuity of \underline{F}_1 , $\underline{F}_1^{-1}(U) \leq y \Leftrightarrow U \leq \underline{F}_1(y)$ (Proposition 1(5) in Embrecht and Hofert (2013)). Hence, we have the following as required.

$$\mathbb{P}(Y_1 \leq y, D = 1) = \mathbb{P}(\underline{F}_1^{-1}(U) \leq y, U < \mathbb{P}(D = 1)) = \mathbb{P}(U \leq \underline{F}_1(y), U < \mathbb{P}(D = 1)) = \underline{F}_1(y).$$

Finally, we need to verify that Y_1 and Y_0 do indeed have the announced distributions. We shall show that the cdf of Y_1 is indeed F_1 . Reasoning as above, we have the following.

$$\begin{aligned} \mathbb{P}(Y_1 \leq y, D = 0) &= \mathbb{P}((F_1 - \underline{F}_1)^{-1}(U - \mathbb{P}(D = 1)) \leq y, U > \mathbb{P}(D = 1)) \\ &= \mathbb{P}(U \leq (F_1 - \underline{F}_1)(y) + \mathbb{P}(D = 1), U > \mathbb{P}(D = 1)) \\ &= (F_1 - \underline{F}_1)(y). \end{aligned}$$

Therefore $\mathbb{P}(Y_1 \leq y) = \mathbb{P}(Y_1 \leq y, D = 1) + \mathbb{P}(Y_1 \leq y, D = 0) = \underline{F}_1(y) + (F_1 - \underline{F}_1)(y) = F_1(y)$ as required.

B.2. Functional features of potential outcomes.

Proof of Proposition 4. Let $0 < q_1 < q_2 < 1$ and let y_1 and y_2 be the q_1 and q_2 quantiles of the distribution of outcomes in Sector d for the self-selected economy. The following holds.

$$\mathbb{P}(Y \leq y_2 | D = d) - \mathbb{P}(Y \leq y_1 | D = d) = \frac{1}{\mathbb{P}(D = d)} (\underline{F}_d(y_2) - \underline{F}_d(y_1)) \geq \underline{F}_d(y_2) - \underline{F}_d(y_1).$$

Hence, $q_2 - q_1 \geq \underline{F}_d(y_2) - \underline{F}_d(y_1)$. In addition, for any $y \in \mathbb{R}$,

$$\mathbb{P}(Y \leq y | D = d) = \frac{\underline{F}_d(y)}{\mathbb{P}(D = d)} = \underline{F}_d(y) + P(D = 1 - d) \frac{\underline{F}_d(y)}{\mathbb{P}(D = d)} \leq \bar{F}_d(y).$$

Finally, under the stochastic dominance condition, $F(y_j) = \mathbb{P}(Y \leq y_j) \leq \mathbb{P}(Y \leq y_j | D = d)$ for $j = 0, 1$. Since the sharp upper bound for the (q_1, q_2) -interquantile range of the distribution of Y_d is given by

$$\text{I}\bar{\text{Q}}\text{R}(q_1, q_2, F_d) = \max \left\{ y_2 - y_1 \left| \begin{array}{l} F(y_1) \leq q_1 \leq \bar{F}_d(y_1), \\ F(y_2) \leq q_2 \leq \bar{F}_d(y_2), \\ q_2 - q_1 \geq \underline{F}_d(y_2) - \underline{F}_d(y_1), \end{array} \right. \right\}$$

these three inequalities imply that the interquantile range $y_2 - y_1$ satisfies the sharp bounds on the interquantile range for the distribution of Y_d .

However, if we relax the first order stochastic dominance condition, we now show that there exist situations, where the interquantile range in Sector d under self-selection is strictly larger than the upper bound for the corresponding interquantile range of the distribution of potential outcomes under random assignment. Let $\mathbb{P}(Y \leq y)$ be continuous. Let y_{02} be defined by $\mathbb{P}(Y \leq y_{02}, D = d) + \mathbb{P}(D = 1 - d) = q_2$. Finally, suppose that $D = 1 - d \Rightarrow Y \leq y_{02}$, so that $\mathbb{P}(Y \leq y)$ coincides with $\mathbb{P}(Y \leq y, D = d) + \mathbb{P}(D = 1 - d)$ on the right of y_{02} . Then the upper bound for the interquantile range of Y_d is $\underline{F}_d(q_2 - \mathbb{P}(D = d)) - \underline{F}_d(q_1 - \mathbb{P}(D = d))$, which can be made lower than the interquantile range for $Y | D = d$, namely $\mathbb{P}(D = d)(\underline{F}_d(q_2) - \underline{F}_d(q_1))$, with a suitable choice of slope for $\mathbb{P}(Y \leq y, D = d)$.

REFERENCES

- AAKVIK, A., J. HECKMAN, and E. VYTLACIL (2005): “Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs,” *Journal of Econometrics*, 125, 15–51.
- ABBING, J., and J. HECKMAN (2007): “Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation,” in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6, chap. 72, pp. 5145–5303. Elsevier: North Holland.
- ALTONJI, J., P. ARCIDIACONO, and A. MAUREL (2016): “The analysis of field choice in college and graduate schools: determinants and wage effects,” *Handbook of the Economics of Education*, 5, 305–396.
- ALTONJI, J., E. BLOM, and C. MEGHIR (2012): “Heterogeneity in human capital investments: High school curriculum, college major, and careers,” *Annual Review of Economics*, 4, 185–223.
- ALTONJI, J., T. ELDER, and C. TABER (2005): “Selection on observed and unobserved variables: assessing the effectiveness of catholic schools,” *Journal of Political Economy*, 113, 151–184.
- ANDREWS, D., and X. SHI (2014): “Nonparametric inference based on conditional moment inequalities,” *Journal of Econometrics*, 179, 31–45.

- ARCIDIACONO, P. (2004): “Ability sorting and the returns to college major,” *Journal of Econometrics*, 121, 343–375.
- BAILLET, F., A. FRANKEN, and A. WEBER (2017): “DZHW Graduate Panel 2009: Data and Methods Report on the Graduate Panel 2009 (1st and 2nd Survey Waves),” Discussion paper, German Centre for Higher Education Research and Science Studies.
- BALKE, A., and J. PEARL (1997): “Bounds on treatment effects from studies with imperfect compliance,” *Journal of the American Statistical Association*, 92, 1171–1176.
- BASU, A., and J. GHOSH (1978): “Identifiability of the multinormal and other distributions under competing risks model,” *Journal of Multivariate Analysis*, 8, 413–429.
- BAYER, P., S. KHAN, and C. TIMMINS (2011): “Nonparametric identification and estimation in a Roy model with common nonpecuniary returns,” *Journal of Business and Economic Statistics*, 29, 201–215.
- BEDFORD, T., and I. MEILIJON (1997): “A Characterization of marginal distributions of (possibly dependent) lifetime variables which right censor each other,” *Annals of Statistics*, 25, 1622–1645.
- BEEDE, D., T. JULIAN, D. LANGDON, G. MCKITTRICK, B. KHAN, and M. DOMS (2011): “Women in STEM: a gender gap to innovation,” Department of Commerce, Economics and Statistics Administration.
- BEFFY, M., D. FOUGÈRE, and A. MAUREL (2012): “Choosing the field of study in postsecondary education: Do expected earnings matter?,” *Review of Economics and Statistics*, 94(1), 334–347.
- BERESTEANU, A., I. MOLCHANOV, and F. MOLINARI (2011): “Sharp identification regions in models with convex predictions,” *Econometrica*, 79, 1785–1821.
- BERESTEANU, A., I. MOLCHANOV, and F. MOLINARI (2012): “Partial identification using random set theory,” *Journal of Econometrics*, 166, 17–32.
- 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, 323–336.
- BORJAS, G. (1987): “Self-selection and the earnings of immigrants,” *American Economic Review*, 115, 531–553.
- BOUDARBAT, B., and C. MONTMARQUETTE (2007): “Choice of Fields of Study of Canadian University Graduates: The Role of Gender and Their Parents’ Education,” Unpublished manuscript.
- BRYNJOLFSSON, E., and A. MCAFEE (2012): *Race against the machine: how the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Digital Frontier Press.
- CAPLIN, A., and B. NALEBUFF (1991): “Aggregation and social choice: a mean voter theorem,” *Econometrica*, 59, 1–23.
- CARD, D. (2001): “Estimating the return to schooling: progress on some persistent econometric problems,” *Econometrica*, 69, 1127–1160.

- CARD, D., and A. PAYNE (2017): “High school choices and the gender gap in STEM,” NBER Working Paper No. 23769.
- CARNEIRO, P., K. HANSEN, and J. HECKMAN (2002): “Removing the veil of ignorance in assessing the distributional impacts of social policies,” Working Paper 8840, NBER.
- CHANDRA, A., and D. STAIGER (2007): “Productivity spillovers in health care: evidence from the treatment of heart attacks,” *Journal of Political Economy*, 115, 103–140.
- CHERNOZHUKOV, V., S. LEE, and A. ROSEN (2013): “Inference on intersection bounds,” *Econometrica*, 81, 667–737.
- CHESHER, A. (2010): “Instrumental variable models for discrete outcomes,” *Econometrica*, 78, 575–601.
- CHESHER, A., A. ROSEN, and K. SMOLINSKI (2013): “An instrumental variable model of multiple discrete choice,” *Quantitative Economics*, 4, 157–196.
- CHETVERIKOV, D. (2013): “Testing regression monotonicity in econometric models,” unpublished manuscript.
- CHEVALIER, A. (2011): “Subject choice and earnings of UK graduates,” *Economics of Education Review*, 30(6), 1187–1201.
- CHIBURIS, R. (2010): “Semiparametric bounds on treatment effect,” *Journal of Econometrics*, 159, 267–275.
- COX, D. (1972): “Regression models and life tables,” *Journal of the Royal Statistical Society, Series B*, 34, 187–220.
- CROWDER, M. (1991): “On the identifiability crisis in competing risks analysis,” *Scandinavian Journal of Statistics*, 18, 223–233.
- CUNHA, F., J. HECKMAN, and S. NAVARRO (2006): “Counterfactual analysis of inequality and social mobility,” in *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, ed. by G. S. F. Stephen L. Morgan, David B. Grusky. Stanford University Press: Palo Alto.
- CUNHA, F., J. HECKMAN, and S. SCHENNACH (2010): “Estimating the technology of cognitive and noncognitive skill formation,” *Econometrica*, 78, 883–931.
- DAHL, G. (2002): “Mobility and the return to education: testing a Roy model with multiple markets,” *Econometrica*, 70, 2367–2420.
- DAYMONT, T. N., and P. J. ANDRISANI (1984): “Job preferences, college major, and the gender gap in earnings,” *Journal of Human Resources*, pp. 408–428.
- DELGADO, M., and J. C. ESCANCIANO (2012): “Distribution-free tests of stochastic monotonicity,” *Journal of Econometrics*, 170, 68–75.
- D’HAULTFOEUILLE, X., and A. MAUREL (2013): “Another look at the identification at infinity of sample selection models,” *Econometric Theory*, 29, 213–224.
- EISENHAUER, P., J. HECKMAN, and E. VYTLACIL (2015): “The generalized Roy model and the cost-benefit analysis of social programs,” *Journal of Political Economy*, 123, 413–443.

- EMBRECHT, P., and M. HOFERT (2013): “A note on generalized inverses,” *Mathematical Methods of Operations Research*, 77, 423–432.
- FAN, Y., and S.-S. PARK (2010): “Sharp Bounds on the Distribution of Treatment Effects and Their Statistical Inference,” *Econometric Theory*, 26, 931–951.
- FAN, Y., and J. WU (2010): “Partial identification of the distribution of treatment effects in switching regime models and its confidence sets,” *Review of Economic Studies*, 77, 1002–1041.
- FIRPO, S., and G. RIDDER (2008): “Bounds on functionals of the distribution of treatment effects,” IEPR Working Paper 08.09.
- FRENCH, E., and C. TABER (2011): “Identification of models of the labor market,” in *Handbook of Labor Economics*, vol. 4, pp. 537–617.
- GALICHON, A., and M. HENRY (2006): “Inference in incomplete models,” unpublished manuscript.
- GALICHON, A., and M. HENRY (2011): “Set identification in models with multiple equilibria,” *Review of Economic Studies*, 78, 1264–1298.
- GRONAU, R. (1974): “Wage comparisons: a selectivity bias,” *Journal of Political Economy*, 82, 1119–1143.
- HECKMAN, J. (1974): “Shadow prices, market wages and labor supply,” *Econometrica*, 42, 679–694.
- HECKMAN, J. (1990): “Varieties of selection bias,” *American Economic Review*, 80, 313–318.
- HECKMAN, J. (1992): “Randomization and social programs,” in *Evaluating Welfare and Training Programs*, ed. by C. Manski, and R. Garfinkle. Harvard University Press: Cambridge.
- HECKMAN, J. (2001): “Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture,” *Journal of Political Economy*, 109, 673–748.
- HECKMAN, J. (2010): “Building bridges between structural and program evaluation approaches to evaluating policies,” *Journal of Economic Literature*, 48, 356–398.
- HECKMAN, J., and B. HONORÉ (1989): “The identifiability of the competing risks model,” *Biometrika*, 76, 325–330.
- HECKMAN, J., and B. HONORÉ (1990): “The empirical content of the Roy model,” *Econometrica*, 58, 1121–1149.
- HECKMAN, J., and G. SEDLACEK (1985): “Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market,” *Journal of Political Economy*, 93, 1077–1125.
- HECKMAN, J., and G. SEDLACEK (1990): “Self-selection and the distribution of hourly wages,” *Journal of Labor Economics*, 8, 329–363.
- HECKMAN, J., J. SMITH, and N. CLEMENTS (1997): “Making the most out of programme evaluation and social experiments: accounting for heterogeneity in programme impacts,” *Review of Economic Studies*, 64, 487–535.
- HECKMAN, J., and C. TABER (2008): “Roy Model,” in *New Palgrave Dictionary of Economics*. Palgrave Macmillan.

- HECKMAN, J., and E. VYTLACIL (1999): “Local instrumental variables and latent variable models for identifying and bounding treatment effects,” *Proceedings of the National Academy of Sciences*, 96, 4730–4734.
- HECKMAN, J., and E. VYTLACIL (2001): “Instrumental variables, selection models and tight bounds on the average treatment effect,” in *Econometric Evaluations of Labour Market Policies*, Lechner, M., and F. Pfeiffer, eds., pp. 1–16. Heidelberg: Springer-Verlag.
- HECKMAN, J., and E. VYTLACIL (2007): “Econometric evaluation of social programs, Part I: Causal models, structural models and econometric policy evaluation,” in *Handbook of Econometrics*, vol. 6B. James Heckman and Edward Leamer, eds.
- HSU, Y.-C., C.-A. LIU, and X. SHI (2016): “Testing generalized regression monotonicity,” IEAS working paper 16-A009.
- HUNT, J., J.-P. GARANT, H. HERMAN, and D. MUNROE (2013): “Why are women underrepresented amongst patentees?,” *Research Policy*, 42, 831–843.
- JAMES, E., N. ALSALAM, J. CONATY, and D.-L. TO (1989): “College quality and future earnings: where should you send your child to college,” *American Economic Review*, 79, 247–252.
- JOVANOVIĆ, B. (1989): “Observable implications of models with multiple equilibria,” *Econometrica*, 57, 1431–1437.
- JUN, S., J. PINKSE, and H. XU (2010): “Tighter bounds in triangular systems,” *Journal of Econometrics*, 161, 122–128.
- KAHN, S., and D. GINTHER (2017): “Women and STEM,” NBER Working Paper No. 23525.
- KÉDAGNI, D., and I. MOURIFIÉ (2016): “Generalized Instrumental Inequalities: Testing IV independence assumption,” unpublished manuscript.
- KELLY, E., P. J. O’CONNELL, and E. SMYTH (2010): “The economic returns to field of study and competencies among higher education graduates in Ireland,” *Economics of Education Review*, 29(4), 650–657.
- KENNY, L. W., L.-F. LEE, G. S. MADDALA, and R. P. TROST (1979): “Returns to College Education: An Investigation of Self-Selection Bias Based on the Project Talent Data,” *International Economic Review*, 20, 775–789.
- KHAN, S., and E. TAMER (2007): “Partial Rank Estimation of Duration Models with General Forms of Censoring,” *Journal of Econometrics*, 23, 309–347.
- KIRKEBØEN, L., E. LEUVEN, and M. MOGSTAD (2016): “Field of study, earnings and self-selection,” *Quarterly Journal of Economics*, 131, 1057–1111.
- KOOPMANS, T., and O. REIERSOL (1950): “The identification of structural characteristics,” *Annals of Mathematical Statistics*, 21, 165–181.
- LEE, L.-F. (1978): “Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables,” *International Economic Review*, 19, 415–433.

- LEE, S., and A. LEWBEL (2013): “Nonparametric identification of accelerated failure time competing risks models,” *Econometric Theory*, 29, 905–919.
- LEE, S., O. LINTON, and Y.-J. WHANG (2009): “Testing stochastic monotonicity,” *Econometrica*, 77, 585–602.
- LEMIEUX, T. (1998): “Estimating the effects of unions on wage inequality in a panel data model with comparative advantage and non-random selection,” *Journal of Labor Economics*, 16, 261–291.
- LEMIEUX, T. (2015): “Occupations, fields of study and returns to education,” *Canadian Journal of Economics*, 14, 1–31.
- MANDELBROT, B. (1962): “Paretian Distributions and Income Maximization,” *Quarterly Journal of Economics*, 76, 57–85.
- MANSKI, C. (2007): *Identification for prediction and decision*. Harvard University Press: Cambridge.
- MANSKI, C., and J. PEPPER (2000): “Monotone instrumental variables: with an application to the returns to schooling,” *Econometrica*, 68, 997–1010.
- MASELLI, I., and M. BEBLAVÝ (2014): “Why so few students do maths and science.,” CEPS Policy Brief No. 313, 10 January 2014.
- MOURIFIÉ, I. (2011): “Sharp bounds on treatment effects,” forthcoming in the *Journal of Econometrics*.
- MULLIGAN, C., and Y. RUBINSTEIN (2004): “The closing of the gender gap as a Roy model illusion,” NBER Working Paper 10892.
- PAREY, M., J. RUHOSE, F. WALDINGER, and N. NETZ (2017): “The selection of high skilled migrants,” *Review of Economics and Statistics*, 99, 776–792.
- PETERSON, A. (1976): “Bounds on a joint distribution function with fixed sub-distribution functions: application to competing risks,” *Proceedings of the National Academy of Sciences*, 73, 11–13.
- ROY, A. (1951): “Some thoughts on the distribution of earnings,” *Oxford Economic Papers*, 3, 135–146.
- SHAIKH, A., and E. VYTLACIL (2005): “Threshold crossing models and bounds on treatment effects: a nonparametric analysis,” NBER Technical Working Paper 307.
- SHAIKH, A., and E. VYTLACIL (2011): “Partial identification in triangular systems of equations with binary dependent variables,” *Econometrica*, 79, 949–955.
- SHAKED, M., and G. SHANTHIKUMAR (2007): *Stochastic Orders*. Springer.
- STOYE, J. (2010): “Partial identification of spread parameters,” *Quantitative Economics*, 1, 323–357.
- TSIATIS, A. (1975): “A nonidentifiability aspect of the problem of competing risks,” *Proceedings of the National Academy of Sciences*, 72, 20–22.
- VASQUEZ-ALVAREZ, R., B. MELENBERG, and A. VAN SOEST (2003): “Selection bias and measures of inequality,” ISSC WP 2003-04.

- WILLIS, R., and S. ROSEN (1979): "Education and self-selection," *Journal of Political Economy*, 87, 507–526.
- ZAFAR, B. (2013): "College major choice and the gender gap," *Journal of Human Resources*, 48(3), 545–595.

UNIVERSITY OF TORONTO

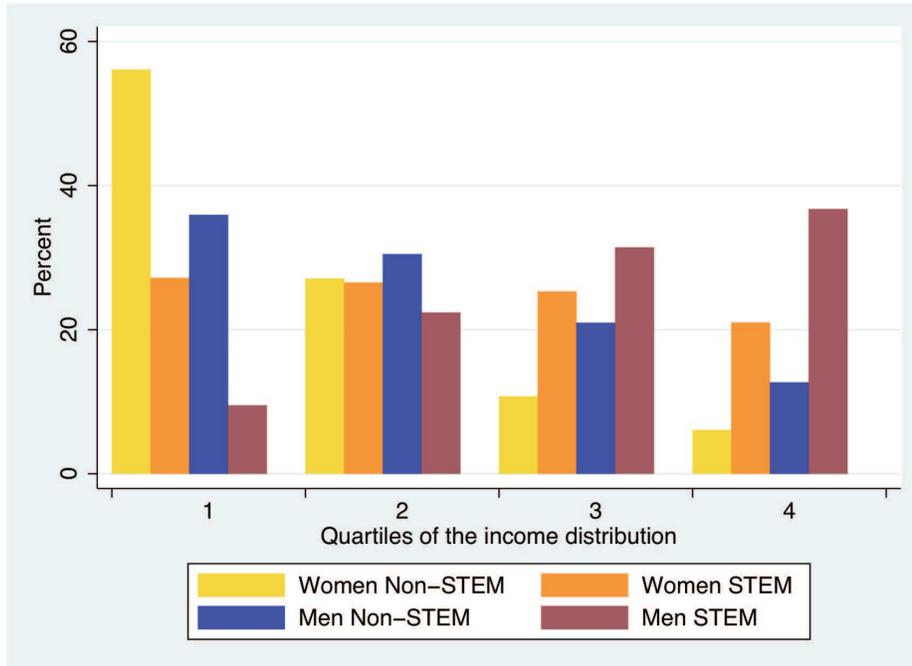
THE PENNSYLVANIA STATE UNIVERSITY

IFO INSTITUTE, MUNICH

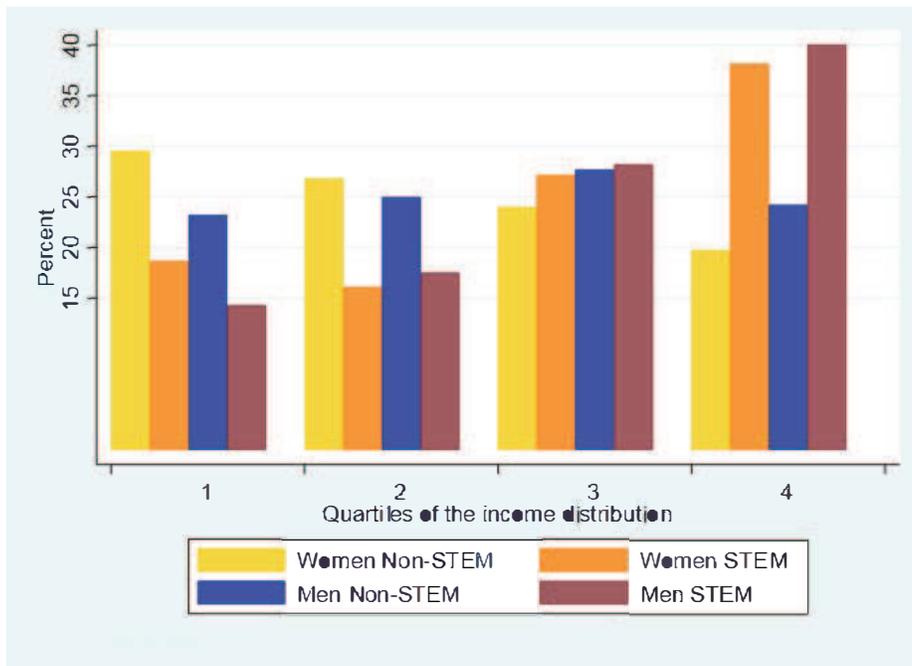
CANADA				
		STEM	OTHER	TOTAL
WOMEN				
	MINORITY	151	841	992
	WHITE	298	4,328	4,626
	TOTAL	449	5,169	5,618
MEN				
	MINORITY	445	528	983
	WHITE	1,052	2,265	3,317
	TOTAL	1,507	2,793	4,300
TOTAL		1,956	7,962	9,918

GERMANY				
		STEM	OTHER	TOTAL
WOMEN				
	MINORITY	243	263	506
	WHITE	1,398	2,542	3,940
	TOTAL	1,641	2,805	4,446
MEN				
	MINORITY	276	74	350
	WHITE	2,170	751	2,921
	TOTAL	2,446	825	3,271
TOTAL		4,087	3,630	7,717

Table 2. Major Choice



(a) Germany



(b) Canada

Figure 3. Quartiles of income distributions by gender and major choice.

CANADA				
		PERMANENT	OTHER	TOTAL
WOMEN				
	STEM	369	79	448
	OTHER	4,009	995	5,004
	TOTAL	4,378	1,074	5,452
MEN				
	STEM	1,332	135	1,467
	OTHER	2,168	448	2,616
	TOTAL	3,500	583	4,083
TOTAL		7,878	1,657	9,535

GERMANY				
		PERMANENT	OTHER	TOTAL
WOMEN				
	STEM	443	804	1,247
	OTHER	499	1702	2,201
	TOTAL	942	2,506	3,448
MEN				
	STEM	769	1,100	1,869
	OTHER	153	508	661
	TOTAL	922	1,608	2,530
TOTAL		1,864	4,114	5,978

Table 3. Permanent employment 1 year (Germany) and 3 years (Canada) after graduation. Each entry is the number of individuals in that category.

CANADA		
	QUÉBEC	REST
WOMEN		
	MINORITY 1.784 (.507)	1.680 (.552)
	WHITE 1.591 (.611)	1.593 (.601)
MEN		
	MINORITY 1.691 (.538)	1.629 (.575)
	WHITE 1.635 (.560)	1.574 (.618)

GERMANY		
	EAST	WEST
WOMEN		
	MINORITY -	2.102 (.054)
	WHITE 1.842 (.036)	1.793 (.024)
MEN		
	MINORITY -	2.165 (.066)
	WHITE 2.133 (.051)	2.090 (.027)

Table 4. Mean survey responses (with standard deviations in parentheses) to a question on the importance of labor market considerations on major choice, from 0 for “not important” to 3 for “very important” in Canada, and from 0 for “not at all important,” to 4 for “very important” for Germany.

SMIV: MOTHER'S EDUCATION

CANADA		TARGET VARIABLE					
		INCOME		PERMANENT		RELATED	
		QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA
WOMEN	MINORITY						
MEN	WHITE		R**		R**	R**	
	MINORITY						
	WHITE						
GERMANY		TARGET VARIABLE					
		INCOME		PERMANENT		RELATED	
		EAST	WEST	EAST	WEST	EAST	WEST
WOMEN	MINORITY	-		-		-	
	WHITE		R***	R*	R***		
MEN	MINORITY	-		-		-	
	WHITE		R***		R***		

Table 5. Test of Roy with imperfect foresight with mother's education as an instrument. "-" indicates the test was not applied to that category because of low sample size.

SMIV: MOTHER'S EDUCATION

CANADA		TARGET VARIABLE					
		INCOME		LEXICOGRAPHIC		RELATED	
		QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA
WOMEN	MINORITY						
MEN	WHITE		R**	R*	R**	R*	
	MINORITY						
	WHITE						
GERMANY		TARGET VARIABLE					
		INCOME		LEXICOGRAPHIC		RELATED	
		EAST	WEST	EAST	WEST	EAST	WEST
WOMEN	MINORITY	-		-		-	
	WHITE	R*	R***	R*	R***		
MEN	MINORITY	-		-		-	
	WHITE		R***		R***		

Table 6. Test of Roy with perfect foresight with mother's education as an instrument. LEXICOGRAPHIC refers to the vector of variables (PERMANENT, RELATED) ordered lexicographically. "-" indicates the test was not applied to that category because of low sample size.

SMIV: MOTHER'S EDUCATION
AND FEMINIZATION OF STEM

		CANADA					
		INCOME		TARGET VARIABLE PERMANENT		RELATED	
		QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA
WOMEN	MINORITY						
	WHITE					R***	R*

		GERMANY					
		INCOME		TARGET VARIABLE PERMANENT		RELATED	
		EAST	WEST	EAST	WEST	EAST	WEST
WOMEN	MINORITY	-		-		-	
	WHITE	R***	R***	R**	R***		

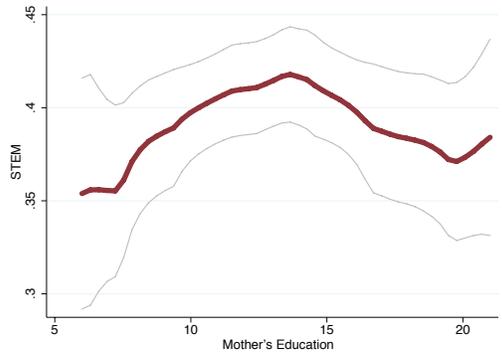
Table 7. Test of Roy with imperfect foresight with mother's education and feminization of STEM faculty as instruments. "-" indicates the test was not applied to that category because of low sample size.

SMIV: MOTHER'S EDUCATION
AND FEMINIZATION OF STEM

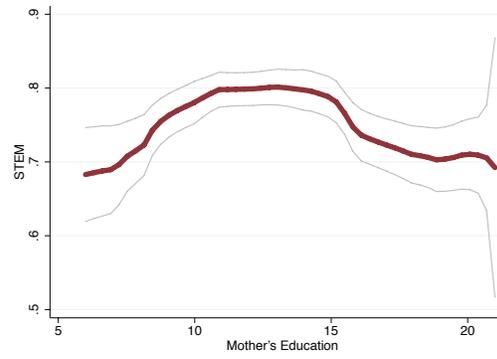
CANADA		TARGET VARIABLE					
		INCOME		LEXICOGRAPHIC		RELATED	
		QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA	QUÉBEC	REST OF CANADA
WOMEN	MINORITY						
	WHITE		R*	R**		R***	R*

GERMANY		TARGET VARIABLE					
		INCOME		LEXICOGRAPHIC		RELATED	
		EAST	WEST	EAST	WEST	EAST	WEST
WOMEN	MINORITY	-		-		-	
	WHITE	R***	R***	R**	R***		

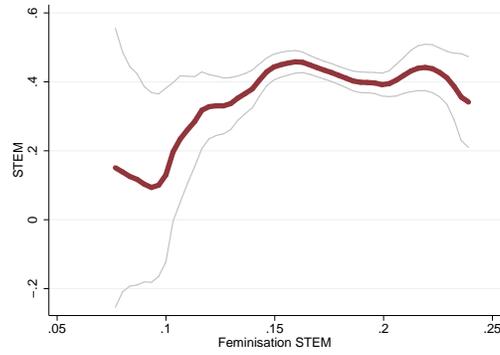
Table 8. Test of Roy with perfect foresight with mother's education and feminization of stem as instruments. LEXICOGRAPHIC refers to the vector of variables (PERMANENT, RELATED) ordered lexicographically. "-" indicates the test was not applied to that category because of low sample size.



(a) Mother's education on women



(b) Mother's education on men



(c) Feminization of STEM on women

Figure 4. Effect of instruments on STEM choices by white residents of the former FRG.

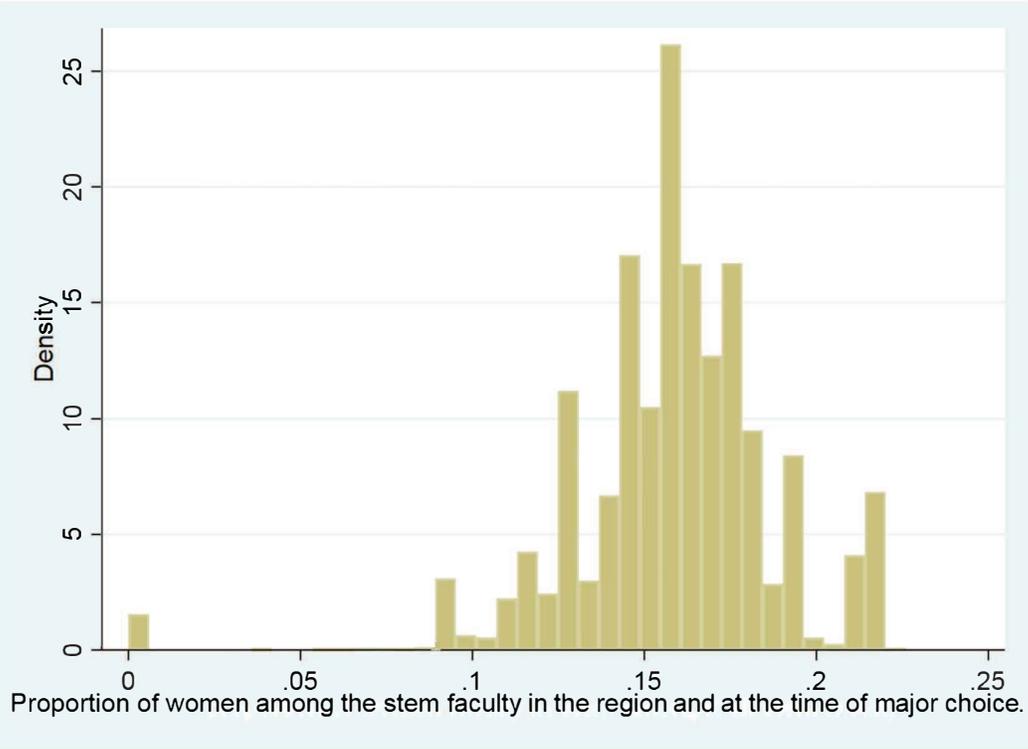
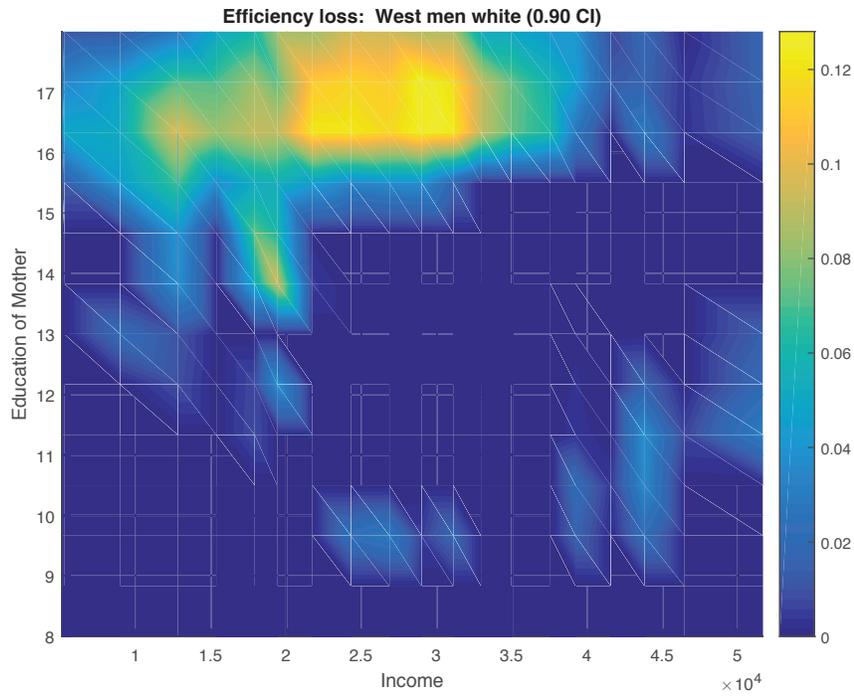
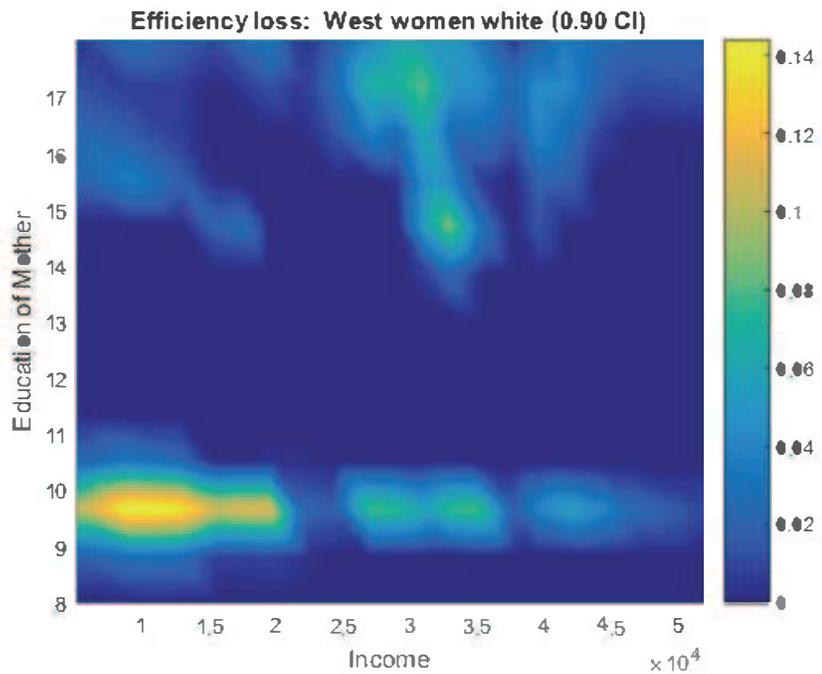


Figure 5. Histogram density of the proportion of women on the STEM faculty in universities in the geographical region of the graduate at the time of major choice.

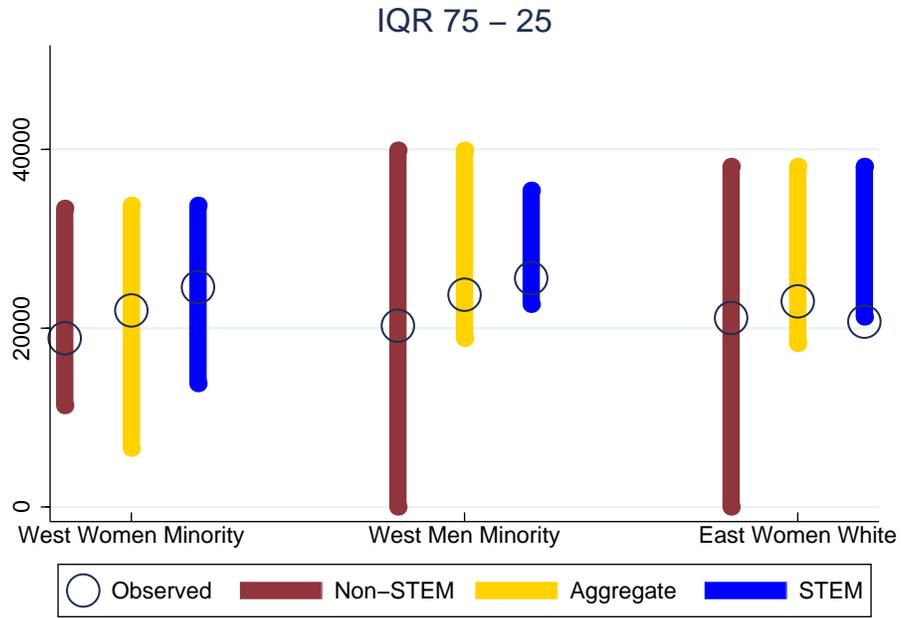


(a) White Men in West Germany

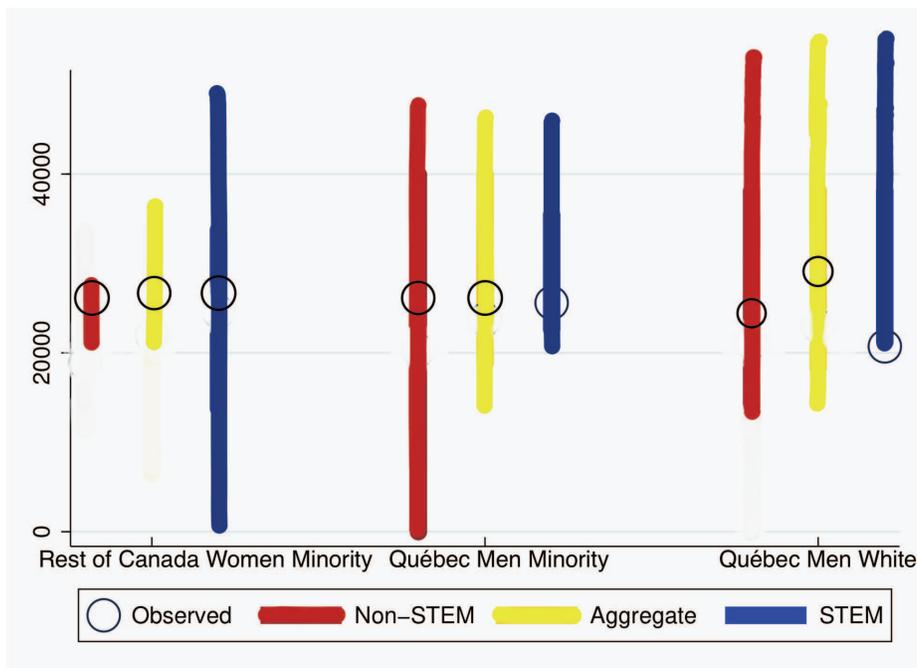


(b) White Women in West Germany

Figure 6. Lower bound of the 90% confidence interval for efficiency loss for white women from the former FRG using mother's education as a SMIV instrument.



(a) Germany



(b) Canada (graph output to be added)

Figure 7. Confidence bands for the interquartile range under the Roy model assumption for categories of individuals, where the latter is not rejected.