Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 15-030

"A Framework for Eliciting, Incorporating, and Disciplining Identification Beliefs in Linear Models"
Third Version

by

Francis J. DiTraglia and Camilo Garcia-Jimeno

A Framework for Eliciting, Incorporating, and Disciplining Identification Beliefs in Linear Models*

Francis J. DiTraglia Camilo Garcia-Jimeno
University of Pennsylvania

This Version: September 9, 2015, First Version: August 29, 2015

Abstract

The identification of causal effects in linear models relies, explicitly and implicitly, on the imposition of researcher beliefs along several dimensions. Assumptions about measurement error, regressor endogeneity, and instrument validity are three key components of any such empirical exercise. Although in practice researchers reason about these three dimensions independently, we show that measurement error, regressor endogeneity and instrument invalidity are mutually constrained by each other and the data in a manner that is only apparent by writing down the full identified set for the model. The nature of this set makes it clear that researcher beliefs over these objects cannot and indeed should not be independent: there are fewer degrees of freedom than parameters. By failing to take this into account, applied researchers both leave money on the table – by failing to incorporate relevant information in estimation – and more importantly risk reasoning to a contradiction by expressing mutually incompatible beliefs. We propose a Bayesian framework to help researchers elicit their beliefs, explicitly incorporate them into estimation and ensure that they are mutually coherent. We illustrate the practical usefulness of our method by applying it to several well-known papers from the empirical microeconomics literature.

Keywords: Partial identification, Beliefs, Instrumental variables, Measurement error, Bayesian econometrics

JEL Codes: C10, C11, C18, C26

*We thank Daron Acemoglu, Siddartha Chib, Philip Dawid, Richard Hahn, Hidehiko Ichimura, Michal Kolesár, Robert McCulloch, Ulrich Müller, and Frank Schorfheide, as well as participants in the Penn Econometrics and Empirical Microeconomics Workshops, and the 2015 NSF-NBER Seminar on Bayesian Inference in Econometrics and Statistics for helpful comments and suggestions.

"Belief is so important! A hundred contradictions might be true."
— Blaise Pascal, *Pensées*

1 Introduction

To identify causal effects from observational data, an applied researcher must augment the data with her beliefs. The exclusion restriction in an instrumental variables (IV) regression, for example, represents the belief that the instrument has no direct effect on the outcome of interest after controlling for other regressors. Although the exclusion restriction can never be directly tested, applied researchers know how to think about it and how to debate it. Indeed because we often have a reasonable idea of which factors make up the regression error term – ability in a wage regression, for example – the exclusion restriction is not simply a matter of opinion but a proposition that we can evaluate in light of other beliefs that we bring to bear on the problem.

In practice, however, not all beliefs are treated equally. The exclusion restriction is what we might call a "formal identification belief," a belief that is directly imposed to achieve identification. In addition to imposing formal beliefs, however, researchers often state a number of other "informal beliefs" that are not imposed in estimation but which may be used, among other things, to interpret results, or reconcile conflicting estimates from different specifications. Papers that report the results of IV regressions, for example, almost invariably state the authors' belief about the sign of the correlation between the endogenous regressor and the error term but fail to exploit this information in estimation.[1] Another commonly stated informal belief involves the extent of measurement error. When researchers uncover an OLS estimate that is substantially smaller than, but has the same sign as its IV counterpart, classical measurement error, with its attendant "least squares attenuation bias," often takes the blame.

In this paper we point out that relegating informal beliefs to second-class status is both wasteful of information and potentially dangerous: beliefs along different dimensions are mutually constrained by each other, by the model, and by the data. This is a simple point which we believe has important implications for applied work but

[1]Referring more than 60 papers published in the top three empirical journals between 2002 and 2005, Moon and Schorfheide (2009) note that "in almost all of the papers the authors explicitly stated their beliefs about the sign of the correlation between the endogenous regressor and the error term; yet none of the authors exploited the resulting inequality moment condition in their estimation."

has been largely overlooked by the literature. By failing to explicitly incorporate relevant information, which they, nevertheless, state and argue about informally, applied researchers leave money on the table and, more importantly, risk reasoning to a contradiction by expressing mutually incompatible beliefs. This intuition is straightforward, but quantifying the nature of the interrelation between these different beliefs and the potential for contradiction is not. We develop a framework to address this general point in the context of a workhorse linear model, and illustrate its practical usefulness. Our starting point is a model of the form

$$
\begin{aligned}
y &= \beta T^* + \mathbf{x}'\gamma + u &(1)\\
T^* &= \pi z + \mathbf{x}'\delta + v &(2)\\
T &= T^* + w &(3)
\end{aligned}
$$

where T^* is a treatment variable, y is an outcome of interest, $\mathbf{x}$ is a vector of exogenous controls, and z is a proposed instrument for T^*. Our goal is to estimate the causal effect of T^* on y, namely β. Unfortunately we observe not T^* but a noisy measure T polluted by measurement error w. Moreover, the true treatment may itself be endogenous: T^* may be correlated with u. While we are fortunate to have an instrument at our disposal, it may not satisfy the exclusion restriction: z is potentially correlated with u. Such a scenario is common in applied work in microeconomics: endogeneity is the rule rather than the exception in social science, the treatments of greatest interest – e.g. the quality of institutions in development economics – are often the hardest to measure, and the validity of a proposed instrument is almost always debatable.

In this draft we assume that T^* is continuous and w is classical measurement error, uncorrelated with all variables in the system besides T. An extension to the case where T^* is binary is currently in progress.[2] We proceed by characterizing the identified set in terms of model primitives over which researchers have formal and informal beliefs, and using this characterization to construct an inferential tool that combines the information in the data with these beliefs in a coherent and transparent way. This tool reveals any inconsistencies in belief that may be present and thus allows researchers to refine and discipline their beliefs about problem at hand. Although our method employs Bayesian reasoning, it can be implemented in a number of different ways that should make it appealing both to frequentist and Bayesian econometricians.

[2] For a discussion of the difference between these two setups, see Section 5.

The approach we follow here is similar in this respect to Kline and Tamer (2015).

While measurement error, endogenous regressors and invalid instruments have all generated voluminous literatures, to the best of our knowledge this is the first paper to provide a full characterization of the identified set when all three problems may be present simultaneously, and to point out that it contains informative restrictions that have not been previously exploited. In a certain sense the lack of attention from the existing literature is unsurprising: a partial identification analysis based on a model that suffers from so many serious problems seems unlikely to produce particularly informative bounds. Yet, as we will show below, by combining data with credible and relatively weak subject-specific beliefs – beliefs that researchers already discuss informally in their research – one can learn a surprising amount both about the causal effect of interest and the coherence of one's beliefs, or lack thereof. We show that measurement error, regressor endogeneity and instrument invalidity are mutually constrained by each other and the data in a manner that is only apparent by characterizing the *full identified set* for the model. In spite of the fact that a single valid instrument solves both the problem of classical measurement error and regressor endogeneity, for example, it is insufficient to carry out a partial identification exercise that merely relaxes the exclusion restriction: values for the correlation between z and u that seem plausible when viewed in isolation may imply wildly implausible amounts of measurement error in T^* or a selection effect with the opposite of the expected sign.

Elicitation is a key ingredient of our framework: before we can impose researcher beliefs about measurement error, regressor endogeneity, and instrument invalidity, we need a way to express each in intuitive, empirically meaningful terms. To this end, we re-parameterize the problem in terms of scale-free variables with bounded support. We express instrument invalidity in terms of ρ_{uz}, the correlation between z and u, and regressor endogeneity in terms of ρ_{T^*u}, the correlation between T^* and u. In the case of classical measurement error, we express the problem not in terms of the measurement error variance but $\kappa = Var(T^*)/Var(T)$, in essence a signal to noise ratio that is conveniently bounded between zero and one. As we discuss further below, it should be fairly easy to impose meaningful bounds on each of these quantities in a typical applied example, and even something as simple as a sign restriction can turn out to be surprisingly informative.

The addition of researcher beliefs is both unavoidable and potentially extremely helpful when using observational data to study causal effects. Nevertheless, whenever one imposes information beyond what is contained in the data, it is crucial to make clear how this affects the ultimate result. This motivates our use of what is referred to in the statistics literature as a *transparent parameterization*.[3] Following Poirier (1998) and Moon and Schorfheide (2012), we show the problem can be decomposed into a vector of partially-identified *structural* parameters θ, and a vector of identified *reduced form* parameters φ in such a way that inference for the identified set Θ for θ depends on the data only through φ. This decomposition has several advantages. First, since the reduced form parameters are identified, inference for this part of the problem is completely standard. The researcher can effectively "drop in" any procedure that generates posterior draws for φ including, if desired, one that exactly matches the usual large-sample frequentist asymptotic distribution. Second, a transparent parameterization shows us precisely where any identification beliefs we may choose to impose enter the problem: the data rule out certain values of φ, while our beliefs amount to placing restrictions on the *conditional* identified set $\Theta(\varphi)$.

In our setting, the reduced form parameter vector φ simply contains the elements of the variance-covariance matrix Σ of the observables $(y, T, \mathbf{x}, z)$. By manipulating the model restrictions to eliminate all other variables, we express the conditional identified set $\Theta(\varphi)$ as a relationship between regressor endogeneity, ρ_{T^*u}, instrument invalidity, ρ_{uz}, κ, the variable that governs the measurement error process.[4] Crucially, the *dimension* of the identified set is strictly smaller than the number of variables used to describe it. Under classical measurement error, for example, the conditional identified set is a heavily constrained, two-dimensional manifold in three-dimensional $(\rho_{T^*u}, \rho_{uz}, \kappa)$–space. This fact makes it clear how easily the constraints of the model, embedded in this identified set, could contradict prior researcher beliefs. Moreover it suggests that knowledge of the form of the identified set could be used to reconcile and refine these beliefs in a way that would not be possible based on introspection alone.

The final ingredient of our procedure is inference for β. Since β is a simple function of θ and φ, and generating posterior draws for φ is straightforward, the problem reduces to sampling from the conditional identified set. We propose a procedure for sampling *uniformly* on the surface $\Theta(\varphi)$. By imposing sign and interval restrictions on the degree of measurement error, endogeneity, and instrument invalidity, we can add prior information to the problem while remaining uniform on the regions of the identified

[3]See, for example, Gustafson (2015).

[4]When T^* is binary, measurement error is governed by two conditional probabilities: α_0 and α_1. For a discussion of these parameters, see Section 5

set that remain. The resulting posterior draws can be used in a number of ways. One possibility is to carry out posterior inference on the worst-case upper and lower bounds for β, in a manner consistent with the usual frequentist analysis of partially identified models. Another is to *average* over the posterior draws in a fully Bayesian fashion yielding a full posterior for β.

This paper relates to a vast literature on partial identification and measurement error. Two recent related papers are Conley et al. (2012), who propose a Bayesian procedure for inference in settings where the exclusion restriction in IV regressions may not hold exactly, and Nevo and Rosen (2012), who derive bounds for a causal effect in the setting where an endogenous regressor is "more endogenous" than the variable used to instrument it is invalid. Our framework encompasses the methods suggested in these two papers, but is more general in several dimensions. First, we allow for measurement error simultaneously with endogeneity and instrument invalidity. More importantly, the central point of our framework is that although imposing restrictions on specific dimensions of an unidentified problem is always informative, it may be misleading unless one has a way to verify the mutual consistency of all beliefs that enter into the problem. In settings where researchers may be willing to implement the partial identification exercises suggested by these papers, our framework will allow them to make sure the added constraints do not require implausible assumptions on measurement error or the sign of the correlation between the endogeneous treatment and the unobservables, for example. Our paper also contributes to a small but growing literature on the Bayesian analysis of partially-identified models, including Poirier (1998), Gustafson (2005), Moon and Schorfheide (2012), Kitagawa (2012), Richardson et al. (2011), Hahn et al. (2015), and Gustafson (2015).

The remainder of this paper is organized as follows. Section 2 characterizes the identified set for the problem, while Section 3 describes our approach to inference. Section 4 presents three empirical examples illustrating the practical usefulness of our method, and Section 5 concludes with some discussion of extensions currently in progress.

2 The Identified Set

To simplify the notation, suppose either that there are no exogenous control regressors $\mathbf{x}$, including a constant, or equivalently that they have been "projected out." In Section 2.4 we explain why this assumption is innocuous and how to accommodate control regressors in practice.

With this simplification, Equations 1–3 become

$$\begin{aligned} y &= \beta T^* + u && (4) \\ T^* &= \pi z + v && (5) \\ T &= T^* + w && (6) \end{aligned}$$

where we assume, without loss of generality, that all random variables in the system are mean zero or have been de-meaned.[5] Our goal is to learn the parameter β, the causal effect of T^*. In general T^* is unobserved: we only observe a noisy measure T that has been polluted by classical measurement error w. We call (u, v, w, z) the "primitives" of the system. Their covariance matrix is as follows:

$$\Omega = Var \begin{bmatrix} u \\ v \\ w \\ z \end{bmatrix} = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} & 0 & \sigma_{uz} \\ \sigma_{uv} & \sigma_v^2 & 0 & 0 \\ 0 & 0 & \sigma_w^2 & 0 \\ \sigma_{uz} & 0 & 0 & \sigma_z^2 \end{bmatrix} \tag{7}$$

Because w represents classical measurement error, it is uncorrelated with u, v, and z as well as T^*. The parameter σ_{uz} controls the endogeneity of the instrument z: unless $\sigma_{uz} = 0$, z is an invalid instrument. Both σ_{uv} and σ_{uz} are sources of endogeneity for the unobserved regressor T^*. In particular,

$$\sigma_{T^*u} = \sigma_{uv} + \pi \sigma_{uz} \tag{8}$$

which we can derive, along with the rest of the covariance matrix for (y, T, T^*, z), from the fact that

$$\begin{bmatrix} y \\ T \\ T^* \\ z \end{bmatrix} = \begin{bmatrix} 1 & \beta & 0 & \beta\pi \\ 0 & 1 & 1 & \pi \\ 0 & 1 & 0 & \pi \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ z \end{bmatrix} \tag{9}$$

along with the assumptions underlying the covariance matrix Ω of (u, v, w, z).

The system we have just finished describing is not identified: without further re-

[5]Equivalently, we can treat the constant term in the first-stage and main equation as exogenous regressors that have been projected out.

strictions we cannot learn the value of β from any amount of data. In particular, neither the OLS nor IV estimators converge in probability to β, instead they approach

$$\beta_{OLS} = \frac{\sigma_{Ty}}{\sigma_T^2} = \beta \left[\frac{\sigma_{T^*}^2}{\sigma_{T^*}^2 + \sigma_w^2} + \frac{\sigma_{T^*u}}{\sigma_{T^*}^2} \right] \tag{10}$$

and

$$\beta_{IV} = \frac{\sigma_{zy}}{\sigma_{Tz}} = \beta + \frac{\sigma_{uz}}{\sigma_{Tz}} \tag{11}$$

where $\sigma_{T^*}^2$ denotes the variance of the *unobserved* true regressor T^*, which equals $\sigma_T^2 - \sigma_w^2$.

Some quantities in the system, however, are identified. Since we observe (T, y, z), we can learn the entries of the covariance matrix Σ of the observables, defined as

$$\Sigma = \begin{bmatrix} \sigma_T^2 & \sigma_{Ty} & \sigma_{Tz} \\ \sigma_{Ty} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{Tz} & \sigma_{yz} & \sigma_z^2 \end{bmatrix} \tag{12}$$

and, as a consequence, the value of the first stage coefficient π since

$$\pi = \frac{\sigma_{T^*z}}{\sigma_z^2} = \frac{\sigma_{Tz}}{\sigma_z^2} \tag{13}$$

where the fact that $\sigma_{T^*z} = \sigma_{Tz}$ follows from Equations 7 and 9.

Although β is unidentified, the observable covariance matrix Σ, along with constraints on the unobserved covariance matrix Ω of the primitives, does impose restrictions on the unobservables. Combined with even relatively weak researcher-specific prior knowledge, these restrictions can sometimes prove surprisingly informative, as we show below. Before we can do this, however, we need to derive the identified set. To aid in this derivation, we first provide a re-parameterization of the problem that will not only simplify the expressions for the identified set, but express it in terms of quantities that are empirically meaningful and thus practical for eliciting beliefs.

2.1 A Convenient Parameterization

The model introduced in the preceding section contains five non-identified parameters: β, σ_{uv}, σ_{uz}, σ_v^2, and σ_w^2. In spite of this, as we will show below, there are only *two* degrees of freedom: knowledge of any two of the five is sufficient to identify the remaining three.

As such we have a choice of how to represent the identified set. Because our ultimate goal is to elicit and incorporate researcher's beliefs, we adopt three criteria for choosing a parameterization:

1. The parameters should be scale-free.

2. The parameter space should be bounded.

3. The parameters should be meaningful in real applications.

Based on these considerations, we define the identified set in terms of the following quantities:

$$\rho_{uz} = Cor(u, z) \tag{14}$$

$$\rho_{T^*u} = Cor(T^*, u) \tag{15}$$

$$\kappa = \frac{\sigma^2_{T^*}}{\sigma^2_T} = \frac{\sigma^2_{T^*}}{\sigma^2_{T^*} + \sigma^2_w} \tag{16}$$

Note that these parameters are not independent of one another. For example, ρ_{T^*u} depends on both κ and ρ_{uz}. This is precisely the point of our analysis: these three quantities are bound together by the assumptions of the model, which allows us to derive the identified set. The first quantity ρ_{uz} is the correlation between the instrument and the main equation error term u. This measures the *endogeneity* of the instrument: the exclusion restriction in IV estimation, for example, corresponds the degenerate belief that $\rho_{uz} = 0$. When critiquing an instrument, researchers often state a belief about the likely sign of this quantity. The second quantity ρ_{T^*u} is the correlation between the *unobserved* regressor T^* and the main equation error term. This measures the *overall* endogeneity of T^*, taking into account both the effect of σ_{uv} and σ_{uz}. While in practice it would be unusual for a researcher to be able to articulate a belief about ρ_{uv}, as pointed out by Moon and Schorfheide (2009), researchers almost invariably state their belief about the sign the quantity ρ_{T^*u} before undertaking an IV estimation exercise making its elicitation easy in many applications.

The third quantity, κ, is somewhat less familiar. In the simple setting we consider here, with no covariates, κ measures the degree of attenuation bias present in the OLS estimator *in the absence of endogeneity in* T^*. In other words, if $\rho_{T^*u} = 0$ then the

OLS probability limit is κ. Equivalently, since $\sigma_{T^*y} = \sigma_{Ty}$

$$\kappa = \left(\frac{\sigma^2_{T^*}}{\sigma^2_T}\right)\left(\frac{\sigma^2_{yT}}{\sigma^2_{yT^*}}\right) = \left(\frac{\sigma^2_{yT}}{\sigma^2_T\sigma^2_y}\right)\left(\frac{\sigma^2_{T^*}\sigma^2_y}{\sigma^2_{yT^*}}\right) = \frac{\rho^2_{yT}}{\rho^2_{yT^*}} \tag{17}$$

so another way to interpret κ is as the ratio of the *observed* R^2 of the main equation and the *unobserved* R^2 that we would obtain if our regressor had not been polluted with measurement error. A third and more general way to think about κ is in terms of signal and noise. If $\kappa = 1/2$, for example, this means that half of the variation in the observed regressor T is "signal," T^*, and the remainder is noise, w. While the other two interpretations we have provided are specific to the case of no covariates, this third interpretation is not.

There are several advantages to parameterizing measurement error in terms of κ rather than the measurement error variance σ^2_w. First, κ has bounded support: it takes a value in $(0, 1]$. When $\kappa = 1$, $\sigma^2_w = 0$ so there is no measurement error. The limit as κ approaches zero corresponds to taking σ^2_w to infinity. Second, writing expressions in terms of κ greatly simplifies our calculations. Indeed, as we will see in the next section, the sample data provide simple and informative bounds for κ. Third, and most importantly, we consider it much easier to elicit beliefs about κ than about σ^2_w. We will consider this point in some detail in the empirical examples that we present below. In the section that follows we will solve for ρ_{uz} in terms of ρ_{T^*u}, κ and the observable covariance matrix Σ. First, however, we will derive bounds on these three quantities.

2.2 Bounds on the Non-Identified Parameters

Our parameterization from the preceding section gives us several obvious bounds: $\rho_{T^*u}, \rho_{uz} \in [-1, 1]$ and $\kappa \in (0, 1]$. Yet there are other, less obvious bounds that come from the two covariance matrices: Σ and Ω. To state these additional bounds, we need an expression for σ^2_v, the variation in T^* not attributable to the instrument z, in terms of κ and observables only. To this end, note that the R^2 of the IV first stage, ρ^2_{Tz}, can be expressed as

$$\rho^2_{zT} = \frac{(\pi\sigma_z)^2}{\sigma^2_T\sigma^2_z} = \frac{\pi^2\sigma^2_z}{\sigma^2_T}$$

Combining this with the fact that $\sigma_T^2 = \sigma_v^2 + \sigma_w^2 + \pi^2\sigma_z^2$, we have

$$1 = \frac{\sigma_v^2 + \sigma_w^2}{\sigma_T^2} + \rho_{Tz}^2$$

Rearranging and simplifying we find that $\rho_{Tz}^2 = \kappa - \sigma_v^2/\sigma_T^2$ and hence

$$\sigma_v^2 = \sigma_T^2(\kappa - \rho_{Tz}^2) \tag{18}$$

We now proceed to construct an additional bound for κ in terms of the elements of Σ. To begin, since we can express κ as $\rho_{Ty}^2/\rho_{T^*y}^2$ and squared correlations are necessarily less than or equal to one, it follows that $\kappa > \rho_{Ty}^2$. Although typically stated somewhat differently, this bound is well known: in fact it corresponds to the familiar "reverse regression bound" for β.[6] As it happens, however, Σ provides an *additional bound* that may be tighter than $\kappa > \rho_{Ty}^2$. Since σ_v^2 and σ_T^2 are both strictly positive, Equation 18 immediately implies that $\kappa \geq \rho_{Tz}^2$. In other words the R^2 of the IV first-stage provides an *upper bound* for the maximum possible amount of measurement error. Given its simplicity, we doubt that we are the first to notice this additional bound. Nevertheless, to the best of our knowledge, it has not appeared in the literature. Taking the best of these two bounds, we have

$$\max\{\rho_{zT}^2, \rho_{yT}^2\} \leq \kappa \leq 1 \tag{19}$$

Recall that κ is *inversely* related to the measurement error variance σ_w^2: larger values of κ correspond to *less* noise. We see from the bound in Equation 19 that larger values of either the first-stage or OLS R-squared leave *less room* for measurement error. This is important because applied econometricians often argue that their data is subject to large measurement error, explaining large discrepancies between OLS and IV estimates, but we are unaware of any cases in which such belief is confronted with these restrictions.

Before proceeding to solve for the identified set, we derive one further bound from the requirement that Ω – the covariance matrix of the model primitives (u, v, w, z) – be positive definite. At first glance it might appear that this restriction merely ensures that variances are positive and correlations bounded above by one in absolute value. Recall, however, that Equation 7 imposes a considerable degree of structure on Ω. In

[6]To see this, suppose that $\rho_{T^*u} = 0$, and without loss of generality that β is positive. Then Equation 10 gives $\beta_{OLS} = \kappa\beta < \beta$. Multiplying both sides of $\kappa > \rho_{Ty}^2$ by β and rearranging gives $\beta < \beta_{OLS}/\rho_{Ty}^2$, and hence $\beta_{OLS} < \beta < \beta_{OLS}/\rho_{Ty}^2$.

particular, many of its elements are assumed to equal zero. Consider the restriction $|\Omega| > 0$. This implies

$$\sigma_w^2 \left[-\sigma_{uv}^2 \sigma_z + \sigma_v^2 \left(\sigma_u \sigma_z^2 - \sigma_{uz}^2\right)\right] > 0$$

but since $\sigma_w^2 > 0$, this is equivalent to

$$\sigma_z^2 \left(\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2\right) > \sigma_v^2 \sigma_{uz}^2$$

Dividing both sides through by $\sigma_u^2 \sigma_z^2 \sigma_v^2$ and rearranging, we find that

$$\rho_{uv}^2 + \rho_{uz}^2 < 1 \tag{20}$$

In other words (ρ_{uz}, ρ_{uv}) must lie within the unit circle: if one of the correlations is very large in absolute value, the other cannot be. To understand the intuition behind this constraint, recall that since v is the residual from the projection of T^* onto z, it is uncorrelated with z by construction. Now suppose that $\rho_{uz} = 1$. If ρ_{uv} were also equal to one, we would have a contradiction: z and v would be perfectly correlated. The constraint given in Inequality 20 rules this out.

As explained above, we will characterize the identified set in terms of ρ_{T^*u}, ρ_{uz} and κ, eliminating ρ_{uv} from the system. Thus, we need to restate Inequality 20 so that it no longer involves ρ_{uv}. To accomplish this, first write

$$\rho_{T^*u} = \left(\frac{\sigma_v}{\sigma_{T^*}}\right) \rho_{uv} + \left(\frac{\pi \sigma_z}{\sigma_{T^*}}\right) \rho_{uz}$$

and then note that $\sigma_v / \sigma_T^* = \sqrt{1 - \rho_{Tz}^2 / k}$ and $\pi \sigma_z / \sigma_{T^*} = \sqrt{\rho_{Tz}^2}$ using Equation 18 and the definition of κ. Combining,

$$\rho_{T^*u} = \left(\sqrt{1 - \rho_{Tz}^2 / \kappa}\right) \rho_{uv} + \left(\sqrt{\rho_{Tz}^2 / \kappa}\right) \rho_{uz} \tag{21}$$

and solving for ρ_{uv},

$$\rho_{uv} = \frac{\rho_{T^*u} \sqrt{\kappa} - \rho_{uz} \rho_{Tz}}{\sqrt{\kappa - \rho_{Tz}^2}} \tag{22}$$

so we can re-express the constraint from Inequality 20 as

$$\left(\frac{\rho_{T^*u}\sqrt{\kappa} - \rho_{uz}\rho_{Tz}}{\sqrt{\kappa - \rho_{Tz}^2}}\right)^2 + \rho_{uz}^2 < 1 \tag{23}$$

2.3 Solving for the Identified Set

We now provide a characterization of the identified set by solving for ρ_{uz} in terms of ρ_{T^*u}, κ and the observables contained in Σ. Rewriting the Equation 11 (the IV estimator), we have

$$\beta = \frac{\sigma_{zy} - \sigma_{uz}}{\sigma_{zT}} \tag{24}$$

and proceeding similarly for Equation 10 (the OLS estimator),

$$\beta = \frac{\sigma_{Ty} - \sigma_{T^*u}}{\kappa\sigma_T^2} \tag{25}$$

Combining Equations 24 and 25, we have

$$\frac{\sigma_{zy} - \sigma_{uz}}{\sigma_{zT}} = \frac{\sigma_{Ty} - \sigma_{T^*u}}{\kappa\sigma_T^2} \tag{26}$$

Now, using Equations 7 and 9, the variance of y can be expressed as

$$\sigma_y^2 = \sigma_u^2 + \beta\left(2\sigma_{T^*u} + \beta\kappa\sigma_T^2\right)$$

Substituting Equation 24 for β, Equation 25 for $\beta\kappa\sigma_T^2$, and rearranging,

$$\left(\sigma_u^2 - \sigma_y^2\right) + \left(\frac{\sigma_{zy} - \sigma_{uz}}{\sigma_{zT}}\right)\left(\sigma_{T^*u} + \sigma_{Ty}\right) = 0 \tag{27}$$

The next step is to eliminate σ_u from our system of equations. First we substitute

$$\begin{aligned} \sigma_{T^*z} &= \sigma_u\sqrt{\kappa}\sigma_T\rho_{T^*u} \\ \sigma_{uz} &= \sigma_u\sigma_z\rho_{uz} \end{aligned}$$

into Equations 26 and 27, yielding

$$\left(\sigma_u^2 - \sigma_y^2\right) + \left(\frac{\sigma_{zy} - \sigma_u\sigma_z\rho_{uz}}{\sigma_{zT}}\right)\left(\sigma_u\sigma_T\sqrt{\kappa}\rho_{T^*u} + \sigma_{Ty}\right) = 0 \tag{28}$$

and

$$\frac{\sigma_{zy} - \sigma_u \sigma_z \rho_{uz}}{\sigma_{zT}} = \frac{\sigma_{Ty} - \sigma_T \sigma_u \rho_{T^*u}}{\kappa \sigma_T^2} \tag{29}$$

Rearranging Equation 29 and solving for σ_u, we find that

$$\sigma_u = \frac{\sigma_{zT}\sigma_{Ty} - \kappa\sigma_T^2\sigma_{zy}}{\sigma_T\sqrt{\kappa}\sigma_{Tz}\rho_{T^*u} - \sigma_z\kappa\sigma_T^2\rho_{uz}} \tag{30}$$

Since we have stated the problem in terms of *scale-free* structural parameters, namely $(\rho_{uz}, \rho_{T^*u}, \kappa)$, we may assume without loss of generality that $\sigma_T = \sigma_y = \sigma_z = 1$. Even if the raw data do not satisfy this assumption, the identified set for the structural parameters is unchanged. Imposing this normalization, the equation for the identified set becomes

$$\left(\widetilde{\sigma}_u^2 - 1\right) + \left(\frac{\rho_{zy} - \widetilde{\sigma}_u\rho_{uz}}{\rho_{zT}}\right)\left(\widetilde{\sigma}_u\sqrt{\kappa}\rho_{T^*u} + \rho_{Ty}\right) = 0 \tag{31}$$

where

$$\widetilde{\sigma}_u = \frac{\rho_{Tz}\rho_{Ty} - \kappa\rho_{zy}}{\sqrt{\kappa}\rho_{Tz}\rho_{T^*u} - \kappa\rho_{uz}} \tag{32}$$

We use the notation $\widetilde{\sigma}_u$ to indicate that normalizing y to have unit variance *does* change the scale of σ_u. Specifically, $\widetilde{\sigma}_u = \sigma_u/\sigma_y$. This does not introduce any complications because we eliminate $\widetilde{\sigma}_u$ from the system by substituting Equation 32 into Equation 31. Note, however, that when $\sqrt{\kappa}\rho_{uz} = \rho_{T^*u}\rho_{Tz}$, Equation 30 has a singularity.

After eliminating $\widetilde{\sigma}_u$, Equation 31 becomes a quadratic in ρ_{uz} that depends on the structural parameters (ρ_{T^*u}, κ) and the reduced form correlations $(\rho_{Ty}, \rho_{Tz}, \rho_{zy})$. Solving, we find that

$$(\rho_{uz}^+, \rho_{uz}^-) = \left(\frac{\rho_{T^*u}\rho_{Tz}}{\sqrt{\kappa}}\right) \pm \left(\rho_{Ty}\rho_{Tz} - \kappa\rho_{zy}\right)\sqrt{\frac{1 - \rho_{T^*u}^2}{\kappa\left(\kappa - \rho_{Ty}^2\right)}} \tag{33}$$

Notice that the fraction under the square root is *always* positive, so both solutions are always real. This follows because $\rho_{T^*u}^2$ must be between zero and one and, as we showed above, $\kappa > \rho_{Ty}^2$. Although the preceding expression always yields two real solutions, one of these is extraneous as it implies a *negative* value for $\widetilde{\sigma}_u$. To see why this is the case, substitute each solution into the *reciprocal* of Equation 32. We have

$$\begin{aligned}
\widetilde{\sigma}_u^{-1} &= \frac{\sqrt{\kappa}\rho_{Tz}\rho_{T^*u}}{\rho_{Ty}\rho_{Tz}-\kappa\rho_{zy}} - \frac{\kappa}{\rho_{Tz}\rho_{Ty}-\kappa\rho_{zy}}\left[\left(\frac{\rho_{T^*u}\rho_{Tz}}{\sqrt{\kappa}}\right) \pm (\rho_{Ty}\rho_{Tz}-\kappa\rho_{zy})\sqrt{\frac{1-\rho_{T^*u}^2}{\kappa\left(\kappa-\rho_{Ty}^2\right)}}\right] \\
&= \frac{\sqrt{\kappa}\rho_{Tz}\rho_{T^*u}}{\rho_{Ty}\rho_{Tz}-\kappa\rho_{zy}} - \left[\left(\frac{\sqrt{\kappa}\rho_{Tz}\rho_{T^*u}}{\rho_{Ty}\rho_{Tz}-\kappa\rho_{zy}}\right) \pm \sqrt{\frac{\kappa(1-\rho_{T^*u}^2)}{\left(\kappa-\rho_{Ty}^2\right)}}\right] \\
&= \mp\sqrt{\frac{\kappa(1-\rho_{T^*u}^2)}{\left(\kappa-\rho_{Ty}^2\right)}}
\end{aligned}$$

Since the quantity inside the square root is *necessarily* positive given the constraints on correlations and κ, we see that ρ_{uz}^{+} is always extraneous. Thus, the only admissible solution is

$$\rho_{uz} = \left(\frac{\rho_{T^*u}\rho_{Tz}}{\sqrt{\kappa}}\right) - (\rho_{Ty}\rho_{Tz}-\kappa\rho_{zy})\sqrt{\frac{1-\rho_{T^*u}^2}{\kappa\left(\kappa-\rho_{Ty}^2\right)}} \tag{34}$$

Along with Inequalities 19 and 23, and the requirement that correlations be less than one in absolute value, Equation 34 gives a complete and minimal characterization of the identified set of the model based on three quantities over which beliefs can be easily articulated. It also reveals that the identified set only has two degrees of freedom, even though researchers can often express beliefs about instrument validity, treatment endogeneity, and measurement error independently. Given any pair from the vector $(\rho_{uz}, \rho_{T^*u}, \kappa)$ and values for the observed moments $(\sigma_T, \sigma_z, \sigma_y, \rho_{Ty}, \rho_{Tz}, \rho_{yz})$ of the covariance matrix Σ, we can solve for the implied value of β using Equation 24. Specifically,

$$\beta = \frac{\sigma_y}{\sigma_z}\left(\frac{\rho_{yz}-\rho_{uz}\widetilde{\sigma}_u}{\rho_{Tz}}\right) \tag{35}$$

using the fact that $\widetilde{\sigma}_u = \sigma_u/\sigma_y$, where $\widetilde{\sigma}_u$ is the standard deviation of the main equation error term from the normalized system, as given in Equation 30, and σ_u is the standard deviation of the main equation error term from the original system. Notice that ρ_{T^*u} and κ enter Equation 35 through $\widetilde{\sigma}_u$. This fact highlights the central point of our analysis: even though *exact knowledge* of σ_{uz} alone would be sufficient to correct the IV estimator, yielding a consistent estimator of β, stating *beliefs* about this quantity alone does not provide a satisfactory solution to the identification problem. For one, because it depends on the scaling of both z and u, it may be difficult to elicit beliefs about σ_{uz}. Although we can learn σ_z from the data, σ_u can only be estimated if we have resolved the identification problem. In contrast, ρ_{uz}, our preferred parameterization, is scale-free. More importantly, however, the form of the identified set makes it clear that our beliefs about ρ_{uz} are *constrained* by any beliefs we may have about ρ_{T^*u} and κ. This observation has two important consequences. First, it provides us with the opportunity to *incorporate* our beliefs about measurement error and the endogeneity of the regressor to improve our estimates. Failing to use this information is like leaving money on the table. Second, it *disciplines* our beliefs to prevent us from reasoning to a contradiction. Without knowledge of the form of the identified set, applied researchers could easily state beliefs that are mutually incompatible *without realizing it*. Our analysis provides a tool for them to realize this and adjust their beliefs accordingly. While we have thus far discussed only beliefs about ρ_{uz}, ρ_{T^*u} and κ, one could also work backwards from beliefs about β to see how they constrain the identified set. We explore this possibility in one of our examples below.

2.4 Accommodating Exogenous Controls

At the beginning of Section 2 we assumed either that there were no control regressors or that they had been projected out. Because the control regressors $\mathbf{x}$ from Equations 1 and 2 are *exogenous*, this is innocuous, as we now show. Without loss of generality, suppose that (T^*, z, y) are mean zero or have been demeaned. Let $(\widetilde{T}^*, \widetilde{T}, \widetilde{y}, \widetilde{z})$ denote the residuals from a linear projection of the random variables (T^*, y, z) on $\mathbf{x}$, e.g. $\widetilde{T}^* = T^* - \Sigma_{T^*\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}$ and so on, where Σ_{ab} is shorthand for $Cov(a, b)$. Then, provided that $(T^*, T, \mathbf{x}, y, z)$ satisfy Equations 1–3, it follows that

$$\widetilde{y} = \beta\widetilde{T}^* + u \tag{36}$$

$$\widetilde{T}^* = \pi\widetilde{z} + v \tag{37}$$

$$\widetilde{T} = \widetilde{T}^* + w \tag{38}$$

since $\mathbf{x}$ is uncorrelated with u and w by assumption and uncorrelated with v by construction. The parameters of this transformed system, β and π, are identical to those of the original system, as are the error terms. And because the transformed system contains no covariates, the analysis presented above applies directly.

The only complication involves the structural parameters κ, ρ_{T^*u} and ρ_{uz}. While it makes sense to elicit researcher beliefs *before* projecting out $\mathbf{x}$, the equations for the identified set presented above will involve not these quantities but their analogues for the transformed system, namely

$$\begin{aligned}\widetilde{\kappa} &= Var(\widetilde{T}^*)/Var(\widetilde{T}) \\ \widetilde{\rho}_{\widetilde{T}^*u} &= Cor(\widetilde{T}^*, u) \\ \widetilde{\rho}_{u\widetilde{z}} &= Cor(\widetilde{z}, u)\end{aligned}$$

Both for the purposes of eliciting beliefs and understanding the identified set, we need to be able to transform back and forth the parameters on their original scale and the versions that project out $\mathbf{x}$. For $\widetilde{\kappa}$, we have

$$\widetilde{\kappa} = \frac{\sigma^2_{T^*} - \Sigma_{T\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\Sigma_{\mathbf{x}T}}{\sigma^2_T - \Sigma_{T\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\Sigma_{\mathbf{x}T}} = \frac{\sigma^2_{T^*}(1 - \Sigma_{T\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\Sigma_{\mathbf{x}T}/\sigma^2_{T^*})}{\sigma^2_T(1 - \Sigma_{T\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\Sigma_{\mathbf{x}T}/\sigma^2_T)} = \frac{\kappa - R^2_{T.\mathbf{x}}}{1 - R^2_{T.\mathbf{x}}} \tag{39}$$

where $R^2_{T.\mathbf{x}}$ denotes the population R-squared from a regression of T on $\mathbf{x}$, and we have used the fact that, since w is classical measurement error, $\Sigma_{T^*\mathbf{x}} = \Sigma_{T\mathbf{x}}$. Equation 39 relates $\kappa \equiv \sigma^2_{T^*}/\sigma^2_T$ for the original system to the analogue $\widetilde{\kappa}$, purely in terms of an identified quantity: $R^2_{T.\mathbf{x}}$. Thus, if the user states beliefs over κ, we can easily transform them to the implied beliefs about $\widetilde{\kappa}$ simply by using the R-squared that results from the regression that projects $\mathbf{x}$ out of T. We can proceed similarly for the other parameters that characterize the identified set. For $\widetilde{\rho}_{u\widetilde{z}}$ we have

$$\widetilde{\rho}_{u\widetilde{z}} = \frac{Cov(z - \Sigma_{z\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\mathbf{x}, u)}{\sigma_u\sqrt{\sigma^2_z - \Sigma_{z\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\Sigma_{\mathbf{x}z}}} = \frac{Cov(z, u)}{\sigma_u\sigma_z\sqrt{1 - \Sigma_{z\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\Sigma_{\mathbf{x}z}/\sigma^2_z}} = \frac{\rho_{uz}}{\sqrt{1 - R^2_{z.\mathbf{x}}}} \tag{40}$$

using the fact that $\mathbf{x}$ is assumed to be uncorrelated with u, where $R^2_{z.\mathbf{x}}$ denotes the R-squared from the population regression of z on $\mathbf{x}$. Finally, for $\widetilde{\rho}_{\widetilde{T}^*u}$, we have

$$\widetilde{\rho}_{\widetilde{T}^*u} = \frac{Cov(T^* - \Sigma_{T\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\mathbf{x}, u)}{\sigma_u\sqrt{\sigma_{T^*} - \Sigma_{T\mathbf{x}}\Sigma_{\mathbf{xx}}\Sigma_{\mathbf{x}T}}} = \frac{Cov(T^*, u)}{\sigma_u\sigma_{T^*}\sqrt{1 - \Sigma_{T\mathbf{x}}\Sigma^{-1}_{\mathbf{xx}}\Sigma_{\mathbf{x}T}/\sigma^2_{T^*}}} = \frac{\rho_{T^*u}}{\sqrt{1 - R^2_{T.\mathbf{x}}/\kappa}} \tag{41}$$

using the fact that $\mathbf{x}$ is uncorrelated with u and $\Sigma_{T^*\mathbf{x}} = \Sigma_{T\mathbf{x}}$.[7]

Since we can always reduce a problem with exogenous covariates to one without, and because we can describe the mapping between the parameters that govern the identified set of the original problem and those of the transformed system, we can

[7]Note that Equation 18 also applies for the partialed-out system. This, combined with Equation 39, implies that κ is always strictly less than $R^2_{T.\mathbf{x}}$, and thus, the expression inside the square root in Equation 41 is guaranteed to be positive.

easily accommodate control variables in the framework derived above. In practice, one simply projects out $\mathbf{x}$ before proceeding, using the R-squared values from these preliminary regressions to transform any user-specified restrictions over $\rho_{uz}, \rho_{T^*u}, \kappa$ into the implied restrictions over $\widetilde{\rho}_{u\widetilde{z}}, \widetilde{\rho}_{\widetilde{T}^*u}$ and $\widetilde{\kappa}$.

3 Inference

We now turn our attention to inference. Our approach is is Bayesian in that we proceed by simulating posterior draws, but many of the inferences we draw can be given a frequentist interpretation if preferred. The key is our decomposition of the identified set from above into two pieces: the reduced form parameters Σ, and the structural parameters $\theta = (\kappa, \rho_{uz}, \rho_{T^*u})$. This allows us to proceed in two simple steps. First, we carry out inference for the reduced form parameter Σ. This part of the problem is completely standard: in particular, the usual large-sample equivalence between Bayesian posterior credible intervals and frequentist confidence intervals holds for these parameters. As such, the user of our method is free to employ *any* desired method of generating posterior draws for Σ in our framework. We propose two simple and relatively uncontroversial possibilities below in Section 3.1. Each draw $\Sigma^{(j)}$ from the posterior for the reduced form parameters leads to an identified set $\Theta(\Sigma^{(j)})$ for the reduced form parameters θ. In our second step, we draw uniformly over $\Theta(\Sigma^{(j)})$, possibly subject to user-specified interval or sign restrictions on the elements of θ. Repeating these two steps yields a collection of posterior draws for $\Sigma, \kappa, \rho_{T^*u}$ and ρ_{uz}, from which we can calculate the implied posterior draws for β.

The resulting draws can be used to conduct inference in a number of different ways. One possibility is to carry out inference for the identified set itself, possibly subject to some restriction on θ. In this approach, the uniform distribution over $\Theta(\Sigma^{(j)})$ is used merely as a computational device rather than a statement of prior belief: we average over $\Sigma^{(j)}$ and take *bounds* over Θ. For example, each conditional identified set $\Theta(\Sigma^{(j)})$ implies an identified set $B(\Sigma^{(j)})$ for β. By averaging over $\Sigma^{(j)}$ we can conduct inference for B, for example by constructing a credible set. In large samples such a credible set will exactly coincide with the corresponding frequentist confidence region since Σ is identified. A committed frequentist who wishes to employ our method could limit herself to drawing inferences of this kind, possibly combined with sign or interval restrictions on θ, e.g. by assuming a positive selection effect: $\rho_{T^*u} > 0$. As we will see below, even fairly mild restrictions can be surprisingly informative. One could also use the draws $\Sigma^{(j)}$ to carry out inference for elements of θ, for example by calculating the "probability of a valid instrument" under a set of restrictions. Even when $\Theta(\Sigma^{(j)})$ is not particularly informative about β, it can very easily rule out a wide range of values for κ, ρ_{T^*u} or ρ_{uz}. For example, suppose that we are considering an example in which we very strongly believe that $\rho_{T^*u} < 0$. Subject to this restriction, a given draw $\Theta(\Sigma^{(j)})$ could very easily rule out the possibility that $\rho_{uz} = 0$. Averaging over $\Sigma^{(j)}$, we can calculate the posterior probability that the identified set is compatible with a valid instrument.

A Bayesian, however, may wish to go further by averaging over draws for $(\kappa, \rho_{T^*u}, \rho_{uz})$ as well as those for Σ. The identified set is a two-dimensional manifold of which partial identification bounds for β, at a given value $\Sigma^{(j)}$, consider only the two worst-case points. From this perspective it seems only natural to consider what *fraction* of the points in this set lead to a particular value for β. Accordingly, one could choose to take our uniform samples from $\Theta(\Sigma^{(j)})$ *literally*. This would amount to placing a uniform prior on the conditional identified set. Moon and Schorfheide (2012) employ such a "reference prior" in their example of a two-player entrance game. The data, of course, can never falsify this conditional prior so one must proceed with extreme caution. Moreover, while a uniform distribution may sound like the natural choice for representing prior ignorance, uniformity in $(\kappa, \rho_{T^*u}, \rho_{zu})$-space could very easily imply a highly informative prior in some different parameterization. Nevertheless, as explained above, we believe that there are compelling reasons to parameterize the problem in terms of ρ_{uz}, ρ_{T^*u} and κ: they are scale-free, empirically meaningful quantities about which researchers are naturally inclined to state beliefs. In most situations, however, these beliefs will be fairly vague making it difficult to elicit an informative prior on the identified set. As such, our preferred approach is to split the difference: rather than taking it completely literally, we treat the conditionally uniform prior as a starting point. In one of the examples below, for example, we consider what kind of deviation from uniformity would be necessary to support a particular belief about β.

The remainder of this section gives the computational details that underlie our inference procedure. Section 3.1 proposes two methods of generating draws for $\Sigma^{(j)}$ while Section 3.2 describes an accept-reject algorithm for sampling uniformly from the conditional identified set. In each section we suppress exogenous covariates for simplicity. If covariates are present, we apply the same methods to the *residuals* that result from regressing (T, y, z) on $\mathbf{x}$.

3.1 Posterior Draws for the Reduced Form Parameters

Because all of its elements are identified, inference for Σ is standard and can be carried out in a number of different ways. Here we consider two simple possibilities.

Our first proposal is based on a large-sample approximation. The idea is to condition on T and z, as is standard in instrumental variables regression, and incorporate sampling uncertainty in σ_{Ty} and σ_{zy} only, applying the central limit theorem *exactly* as one does when deriving the frequentist large-sample distribution of IV and OLS estimators. To begin, let

$$\begin{aligned} \varepsilon_T &= (y - E[y]) - \beta_T(T - E[T]) \\ \varepsilon_z &= (y - E[y]) - \beta_z(z - E[z]) \end{aligned}$$

where $\beta_T = \sigma_{Ty}/\sigma_T^2$, and $\beta_z = \sigma_{zy}/\sigma_z^2$. While neither β_T nor β_z equals the true treatment effect β, the parameters of both of these regressions are identified. Under the standard regularity conditions for linear regression, we have

$$\begin{bmatrix} \sqrt{n}\left(\widehat{\beta}_T - \beta_T\right) \\ \sqrt{n}\left(\widehat{\beta}_z - \beta_z\right) \end{bmatrix} \to_d B \begin{bmatrix} M_T \\ M_z \end{bmatrix} \tag{42}$$

where $\widehat{\beta}_T = \widehat{\sigma}_{Ty}/\widehat{\sigma}_T^2$ and $\widehat{\beta}_z = \widehat{\sigma}_{zy}/\widehat{\sigma}_z^2$ are the least-squares estimators of β_T and β_z, $(M_T, M_z)' \sim N(0, V)$, and

$$B = \begin{bmatrix} 1/\sigma_x^2 & 0 \\ 0 & 1/\sigma_z^2 \end{bmatrix},\ V = E\begin{bmatrix} T^2\varepsilon_T^2 & zT\varepsilon_z\varepsilon_T \\ zT\varepsilon_z\varepsilon_T & z^2\varepsilon_z^2 \end{bmatrix}. \tag{43}$$

Note that V depends not on the *structural* error u but on the *reduced form* errors $\varepsilon_T, \varepsilon_z$. By construction ε_T is uncorrelated with T and ε_z is uncorrelated with z but the reduced form errors are *necessarily* correlated with each other. Now, using Equations 42 and 43 we see that

$$\begin{bmatrix} \sqrt{n}\left(\widehat{\sigma}_{Ty} - \sigma_{Ty}\right) \\ \sqrt{n}\left(\widehat{\sigma}_{zy} - \sigma_{zy}\right) \end{bmatrix} \to_d B^{-1}B \begin{bmatrix} M_T \\ M_z \end{bmatrix} = \begin{bmatrix} M_T \\ M_z \end{bmatrix} \tag{44}$$

and thus, in large samples

$$\begin{bmatrix} \widehat{\sigma}_{Ty} \\ \widehat{\sigma}_{zy} \end{bmatrix} \approx N\left(\begin{bmatrix} \sigma_{Ty} \\ \sigma_{zy} \end{bmatrix}, \widehat{V}/n\right) \tag{45}$$

where $\widehat{V}$ is the textbook robust variance matrix estimator, namely

$$\widehat{V} = \frac{1}{n}\sum_{i=1}^{n} \begin{bmatrix} T_i^2 \hat{\varepsilon}_{Ti}^2 & z_i T_i \hat{\varepsilon}_{zi} \hat{\varepsilon}_{Ti} \\ z_i T_i \hat{\varepsilon}_{zi} \hat{\varepsilon}_{Ti} & z_i^2 \hat{\varepsilon}_{zi}^2 \end{bmatrix}$$

where $\widehat{\varepsilon}_{Ti}$ denotes the i^{th} residual from the β_T regression and $\widehat{\varepsilon}_{zi}$ the i^{th} residual from the β_z regression. Since we are working solely with identified parameters, the usual large-sample equivalence between a Bayesian posterior and frequentist sampling distribution holds. Accordingly, we propose to generate draws for σ_{Ty} and σ_{zy} according to

$$\begin{bmatrix} \sigma_{Ty}^{(j)} \\ \sigma_{zy}^{(j)} \end{bmatrix} \sim N\left(\begin{bmatrix} \widehat{\sigma}_{Ty} \\ \widehat{\sigma}_{zy} \end{bmatrix}, \widehat{V}/n\right) \tag{46}$$

Combining these draws with the *fixed* values $\widehat{\sigma}_T^2, \widehat{\sigma}_z^2$ and $\widehat{\sigma}_{zT}$, since we are conditioning on z and T, yields posterior draws for Σ based on a large-sample normal approximation, namely

$$\Sigma^{(j)} = \begin{bmatrix} \widehat{\sigma}_T^2 & \sigma_{Ty}^{(j)} & \widehat{\sigma}_{Tz} \\ \sigma_{Ty}^{(j)} & \widehat{\sigma}_y^2 & \sigma_{zy}^{(j)} \\ \widehat{\sigma}_{Tz} & \sigma_{zy}^{(j)} & \widehat{\sigma}_z^2 \end{bmatrix} \tag{47}$$

Drawing $\Sigma^{(j)}$ based on this large-sample approximation is simple, robust to heteroskedasticity, and likely to appeal broadly. Unfortunately this approach is not guaranteed to produce positive definite draws. When the sample size is small, as in our example from Section 4.1, this can be problematic.

A solution to this problem is to proceed in a fully Bayesian fashion rather than using an approximation based on the Central Limit Theorem. There are many possible ways to accomplish this. One simple possibility is to posit a joint normal likelihood for (T, y, z) and place a Jeffrey's prior on Σ. In particular if,

$$\begin{bmatrix} T_i \\ y_i \\ z_i \end{bmatrix} \sim \text{iid } N_3(\boldsymbol{\mu}, \Sigma), \quad \pi(\boldsymbol{\mu}, \Sigma) \propto |\Sigma|^{-2} \tag{48}$$

then the marginal posterior for Σ is

$$\Sigma|\mathbf{T},\mathbf{y},\mathbf{z} \sim \text{Inverse-Wishart}(n-1, S) \tag{49}$$

where

$$S = \sum_{i=1}^{n} \begin{bmatrix} (T_i - \bar{T}) \\ (y_i - \bar{y}) \\ (z_i - \bar{z}) \end{bmatrix} \begin{bmatrix} (T_i - \bar{T}) & (y_i - \bar{y}) & (z_i - \bar{z}) \end{bmatrix} \tag{50}$$

The Jeffrey's prior approach has several advantages. First, as it involves a non-informative prior, it requires no input from the user. Second, it is guaranteed to produce positive definite draws for Σ which is important given that our derivation of the identified set depends on this condition. Nevertheless, the Jeffrey's prior approach has several disadvantages. For one, the use of a Jeffrey's prior is not without controversy in Bayesian econometrics. But more fundamentally, it may seem odd to model the *joint* distribution of (T, z, y) given that the typical regression problem, Bayesian or frequentist, involves a model of the *conditional* distribution of y given T and z. This may be less of a concern in a examples featuring a large number of exogenous control regressors. Since these are projected out before proceeding, we are in effect positing a normal distribution only for the *residuals* of the regressions of (T, y, z) on $\mathbf{x}$. In the empirical example from Section 4.3, for example, Protestant share, distance to Wittenberg, and literacy rate are highly non-normal. The residuals after projecting out a large number of demographic controls, however, are plausibly Gaussian. Nevertheless, in the examples below, we will only employ the Jeffrey's prior approach in situations where the sample size is too small for our large-sample approximation to ensure positive definite draws for Σ.

3.2 Uniform Draws on the Conditional Identified Set

To generate uniform draws on $\Theta(\Sigma^{(j)})$ we employ an accept-reject algorithm, similar to that proposed by Melfi and Schoier (2004). We proceed in two steps. First we draw $\kappa^{(\ell)} \sim \text{Uniform}(\underline{\kappa}, \overline{\kappa})$ independently of $\rho_{T^*u}^{(\ell)} \sim \text{Uniform}(\underline{\rho}_{T^*u}, \overline{\rho}_{T^*u})$. Absent any prior restrictions that further restrict the support of κ or ρ_{T^*u}, we take $\underline{\kappa} = \max\left\{(\rho_{zT}^2)^{(\ell)}, (\rho_{Ty}^2)^{(\ell)}\right\}$, $\overline{\kappa} = 1$, $\underline{\rho}_{T^*u} = -1$ and $\overline{\rho}_{T^*u} = 1$. We then solve for $\rho_{uz}^{(\ell)}$ via Equation 34 and check whether it lies in the interval $[\underline{\rho}_{uz}, \overline{\rho}_{uz}]$. Absent any prior restrictions on ρ_{uz}, we take this interval to be $[-1, 1]$. If $\rho_{uz}^{(\ell)}$ lies in this region and if the triple $(\rho_{uz}, \rho_{T^*u}, \kappa)$ satisfies Inequality 23, we accept draw ℓ; otherwise we reject it. We repeat this process until we have L draws on the identified set. While these draws are uniform when projected into the (κ, ρ_{T^*u}) plane, they are *not* uniform on the identified set itself. To make them uniform, we need to re-weight each draw based on the local surface area of the identified set at that point. By "local surface area" we refer to the quantity

$$M\left(\rho_{T^*u}, \kappa\right) = \sqrt{1 + \left(\frac{\partial \rho_{uz}}{\partial \rho_{T^*u}}\right)^2 + \left(\frac{\partial \rho_{uz}}{\partial \kappa}\right)^2} \tag{51}$$

which Apostol (1969) calls the "local magnification factor" of a parametric surface. The derivatives required to evaluate the function M are

$$\frac{\partial \rho_{uz}}{\partial \rho_{T^*u}} = \frac{\rho_{Tz}}{\sqrt{\kappa}} + \frac{\rho_{T^*u}\left(\rho_{Ty}\rho_{Tz} - \kappa\rho_{zy}\right)}{\sqrt{\kappa\left(\kappa - \rho_{Ty}^2\right)\left(1 - \rho_{T^*u}^2\right)}} \tag{52}$$

and

$$\frac{\partial \rho_{uz}}{\partial \kappa} = -\frac{\rho_{T^*u}\rho_{Tz}}{2\kappa^{3/2}} + \sqrt{\frac{1 - \rho_{T^*u}^2}{\kappa\left(\kappa - \rho_{Ty}^2\right)}}\left\{\rho_{zy} + \frac{1}{2}\left(\rho_{Ty}\rho_{Tz} - \kappa\rho_{zy}\right)\left[\frac{1}{\kappa} + \frac{1}{\kappa - \rho_{Ty}^2}\right]\right\} \tag{53}$$

To accomplish the re-weighting, we first evaluate $M^{(\ell)} = M(\rho_{T^*u}^{(\ell)}, \kappa^{(\ell)})$ at each draw ℓ that was accepted in the first step. We then calculate $M_{max} = \max_{\ell=1,\ldots,L} M^{(\ell)}$ and *resample* the draws $\left(\rho_{uz}^{(\ell)}, \rho_{T^*u}^{(\ell)}, \kappa^{(\ell)}\right)$ with probability $p^{(\ell)} = M^{(\ell)}/M_{max}$.

4 Empirical Examples

We now consider several examples drawn from the applied literature to illustrate the methods proposed above: the first considers the effect of institutions on income per capita, a second considers the estimation of returns to schooling, and a third explores the causal effect of the Protestant reformation on literacy.

4.1 The Colonial Origins of Comparative Development

We begin by considering the main specification of Acemoglu et al. (2001), who use early settler mortality as an instrument to study the effect of institutions on GDP per capita based on cross-country data for a sample of 64 countries. The main equation is

$$\log \text{GDP/capita} = \text{constant} + \beta\,(\text{Institutions}) + u$$

and the first stage is

$$\text{Institutions} = \text{constant} + \pi\,(\log \text{Settler Mortality}) + v$$

This leads to an OLS estimate of $\widehat{\beta}_{OLS} = 0.52$ (S.E. = 0.06) and an IV estimate that is nearly twice as large, $\widehat{\beta}_{IV} = 0.94$ (S.E. = 0.16), a difference which the authors attribute to measurement error:

> This estimate is highly significant ... and in fact larger than the OLS estimates... This suggests that measurement error in the institutions variables that creates attenuation bias is likely to be more important that reverse causality and omitted variables biases. (p. 1385)

But to what extent can measurement error explain this disparity? Can we use our framework to assess whether measurement error is *more important* than omitted variables? In their paper Acemoglu et al. (2001) state a number of beliefs that are relevant for this exercise. First, their discussion suggests there is likely a positive correlation between "true" institutions and the main equation error term u. For example, they mention that such a correlation could arise through reverse causality if wealthier societies can afford "better" institutions.[8] Second, by way of a footnote (see footnote 19 in their paper) that uses a second measure of institutions as an instrument for the first, they argue that measurement error could be substantial. Translating their calculations from this footnote into our framework implies a value of κ equal to 0.6.[9] This would correspond to 40% of the variation in the observed measure of institutions being noise. Since the 0.6 estimate for κ is subject to sampling uncertainty, we consider a prior that

[8]The authors also discuss several potential omitted variables, such as legal origin and British culture, which are likely to be positively correlated contemporary institutional quality.

[9]The calculation is as follows. Suppose we have two measures of institutions, T_1 and T_2, each subject to classical measurement error: $T_1 = T^* + w_1$ and $T_2 = T^* + w_2$. Because the measurement error is classical, both T_1 and T_1 suffer from *precisely* the same degree of endogeneity, because they inherit this problem from T^* alone. Thus, the probability limit of OLS based on T_1 is $\kappa(\beta + \sigma_{T^*u}/\sigma^2_{T^*})$ while that of IV using T_2 to instrument for T_1 is $\beta + \sigma_{T^*u}/\sigma^2_{T^*}$. Taking the ratio of the two identifies κ. In this particular example $0.52/0.87 \approx 0.6$. The standard error for the numerator is 0.06 while that for the denominator is 0.16.

restricts κ to lie in the interval $[0.45, 0.82]$.[10]

We begin by computing the full identified set for this example evaluated at the MLE for the matrix Σ of reduced form parameters, which we present in Figure 1a. The region in green restricts attention to the two beliefs mentioned above: $\rho_{T^*u} > 0$, and $\kappa \in [0.45, 0.82]$. As the figure illustrates, those quite reasonable beliefs already shrink the identified set considerably, although the data by itself further constrains κ to be at least 0.54 at the MLE for Σ (the constraint in equation 19 binds). The identified set also allows us to locate the region corresponding to $\beta > 0$, i.e. a positive effect of institutions on prosperity. This corresponds to the blue region in Figure 1b. Figure 1c then "zooms in" to the portion of the identified set that corresponds to the prior beliefs, still indicating the points that lead to a positive value of β in blue. Reassuringly from the authors' perspective, less than 4% of the restricted identified set evaluated at the MLE for Σ maps into a negative value for β. Nevertheless, notice that at the MLE for Σ the restricted identified set completely rules out $\rho_{uz} = 0$. Figure 1c shows that under the prior, log settler mortality must be *negatively* correlated with the error. Moreover, this correlation is at least 0.2 in magnitude. We have thus encountered a contradiction in beliefs; evaluated at the most likely values for the reduced form parameters, the model rules out simultaneously believing that the instrument is valid, that 18-55% of the measured variation in institutions is noise, and that institutions are positively correlated with the error.

Of course, examining only the MLE for Σ ignores sampling variability which is likely to be significant in this example as it only includes 64 observations.[11] There are two kinds of probabilities that one can calculate using our framework: those that involve only *bounds* on the identified set, incorporating posterior uncertainty over Σ only, and those that additionally average over the *uniform draws* on the identified set. We begin by considering two probabilities of the first kind. First, while the lower bound for κ at the MLE is 0.54, allowing for sampling variability in Σ yields a range of possible values for this lower bound. The symmetric 90% posterior credible region for the lower bound on κ, imposing *no* prior restrictions on any of the unidentified parameters, is $(0.40, 0.67)$. In other words, the data tell us that it is extremely unlikely that more than 60% of the measured variation in institutions is noise. Second, we

[10]This range is roughly equal to taking upper and lower bounds for κ based on a one standard error interval around 0.52 and 0.87.

[11]For this example, our frequentist large-sample approach is infeasible as the small sample size of 64 observations leads to a substantial number of non-positive definite draws for Σ. Thus, we carry out inference using our fully Bayesian approach with a Jeffery's prior.

can also calculate probabilities that condition on our interval restrictions. Restricting $\rho_{T^*u} > 0$ and $\kappa \in [0.45, 0.82]$ we find that nearly 48% of the posterior density for Σ is *incompatible* with a valid instrument. In other words, each draw $\Sigma^{(j)}$ yields an identified set like the one depicted for $\widehat{\Sigma}_{MLE}$ in Figure 1c and while the shapes of these sets differ, 48% of them exclude $\rho_{uz} = 0$ just as the set for the MLE does.

The preceding two probability statements are informative but also somewhat coarse. To take a particularly stark example, even under our prior support restrictions, 100% of the draws for Σ yield identified sets that allow for the possibility of a negative β. Notice that Figure 1c, evaluated at the MLE, also does not rule out negative values of β. Nevertheless, the corresponding region of the identified set, depicted in green, is a very small fraction of the total area. The same is true for most of the draws for Σ. This motivates a second kind of probability statement, one that averages both over the draws for Σ *and* and over those for $(\rho_{uz}, \rho_{T^*u}, \kappa)$. Unlike the statements about bounds from above, however, inferences constructed in this manner will depend on the assumption of uniformity over the identified set. One need not take this prior literally, however, to learn a great deal from the exercise. For example, we can consider the kinds of deviations from uniformity that would be required to support particular beliefs. Returning to our point about negative values for β, under a conditionally uniform prior for the structural parameters ρ_{T^*u}, ρ_{zu} and κ restricted to $\rho_{T^*u} > 0$ and $\kappa \in [0.45, 0.82]$, just over six percent of the posterior density for β lies below zero, as shown in Figure 2b. This implies that one would need a very strange prior over the conditional identified set, one that disproportionately favors extremely large values of ρ_{T^*u} and extremely negative values of ρ_{uz}, to assign substantial posterior probability to a negative β. We can use the same idea to re-examine our earlier probability statement concerning instrument validity. Averaging over both Σ and the uniform draws on the conditional identified set, we see that the restriction $\rho_{T^*u} > 0$ combined with $\kappa \in [0.45, 0.82]$ gives a posterior probability of 82 percent that ρ_{uz} is less than -0.1. The main implication of this fact is that the IV estimator is likely somewhat of an overestimate in this example, as we see from Figure 2b.[12] The posterior median for β, for example, is around 0.73 compared to an IV point estimate of 0.94.

We learn quite a lot from our framework in this example. The main result of Acemoglu et al. (2001) continues to hold: in spite of the fact that Settler Mortality is

[12] An instrument that is negatively correlated with the structural equation error term will result in a downward-biased estimate precisely when the first-stage coefficient is negative, as it is in this example.

likely negatively correlated with u under reasonable prior beliefs, it appears that the effect of institutions on income per capita is almost certainly positive.[13]

4.2 The Returns to Schooling

In a second example we apply our framework to the classic question in labor economics related to the estimation of the returns to schooling. We use a sample of 741 males from Blackburn and Neumark (1992), which contains labor market information for the year 1980. Following some of the earlier literature on the subject (Griliches (1979), Card (1999)), we propose using father's education as the instrument for the individual's years of schooling. The main equation is

$$\log(\text{wage rate}) = \text{constant} + \beta\,(\text{Years of Schooling}) + \mathbf{x}'\gamma + u$$

and the first stage is

$$\text{Years of Schooling} = \text{constant} + \pi\,(\text{Father's Years of Schooling}) + \mathbf{x}'\delta + v$$

The set of covariates $\mathbf{x}$ we include contains IQ, age, age squared, marital status, race, a South dummy, and an urban dummy. We find an OLS estimate of 0.039 (S.E. = 0.0064), and an IV estimate of 0.072 (S.E. = 0.024). Both of these estimates appear to be small relative to the rest of the literature, and in fact most labor economists would not believe the exclusion restriction behind the choice of father's schooling as an instrument. On the other hand, the literature has pointed out that measurement error in schooling, even if small, may be an important source of attenuation bias in Mincerian regressions. In a well-known study, Ashenfelter and Krueger (1994) use the responses of identical twins about their siblings' education to assess the extent of measurement error in self-reported years of schooling. This allows them to estimate the so-called

[13] This example also illustrates the usefulness of our framework as a tool for *refining* researcher beliefs. Our initial prior restriction for κ was not that it lie in the interval $[0.45, 0.82]$, but rather $\kappa < 0.6$. This belief was communicated to us in personal correspondence by one of the authors of Acemoglu et al. (2001). Based on footnote 19 of the paper, he expressed the belief that at least 40 percent of the measured variation in quality of institutions was likely to be noise. Imposing this belief along with $\rho_{T^*u} > 0$, however, leads to an empty identified set for just under 25 percent of the draws for Σ. In other words, the data strongly refute this prior. This left us with the choice of either relaxing $\rho_{T^*u} > 0$, which would seem strange in this example, or allowing for the possibility of less measurement error. Accordingly, we chose the latter route, leading to the analysis presented above where none of the identified sets are empty.

"reliability ratio" (κ in our notation):

> The two estimates of the reliability ratio for the twins schooling levels ... are 0.92 and 0.88. These estimates indicate that between 8 and 12 percent of the measured variance in schooling levels is error. Previous estimates of the reliability ratio in schooling levels ... have ranged between 0.8 and 0.93 ... (p. 1161)

These estimates are in close agreement to most others in the literature (see Card, 1999). Being based on self-reported years of schooling, we would expect even less measurement error in our sample from Blackburn and Neumark (1992), which uses administrative records. Nevertheless, our prior will restrict κ to take values larger than 0.8, thus allowing for noise to be up to 20% of the total variation in years of schooling. Beliefs regarding the partial correlation between schooling and unobserved ability, ρ_{T^*u} here, are somewhat more contentious among labor economists. Although most seem to agree this correlation should be positive[14] and most likely not larger than 0.8, some may consider the possibility of a weak but negative correlation between schooling and ability (e.g. Ashenfelter and Krueger, 1994; Erickson, 1993).

Figure 3a depicts the identified set at the MLE for Σ. In this case, the data rules out κ's below 0.4 only, and does not rule out any values for either ρ_{uz} or ρ_{T^*u}. The green region depicts our prior for this problem: reliability ratios above 0.8, and a positive but not excessively large correlation between education and unobserved ability: between 0 and 0.8. Figure 3b then depicts in blue the region over the full identified set, evaluated at the MLE, mapping into positive values for β. Notice that given the data, positive returns to schooling are more likely for negative values of either the ability bias or the instrument invalidity. We see from Figures 4a and 4b that, without imposing prior beliefs, the model and the data allow for a huge range for the returns to schooling: as small as -400% or as large as $+400\%$. Figure 3c then imposes our prior, which we believe represents the consensus among labor economists. Restricting the problem to this subset of the parameter space has several implications. First, we learn that believing that schooling and ability are positively correlated rules out a valid instrument at the MLE for Σ. Averaging over the sampling variation, the probability of father's education being a valid instrument is only 7%. That is, 93% of the posterior

[14] Berhman and Rosenzweig (1999) review the literature, and also point out that the negative "selection effect" estimated by Ashenfelter and Krueger (1994) is not robust to the use of a larger sample size.

density for Σ implies an identified set that does not include $\rho_{uz} = 0$. Perhaps more surprisingly, only 20% of the points in the identified set evaluated at the MLE for Σ is consistent with a positive value for β. Figure 4 illustrates this more clearly. Once we introduce our prior, the data and the model make positive returns to schooling very unlikely. In particular only a strong prior, one that assigns an overwhelming amount of density to the region where both ρ_{T^*u} and ρ_{zu} are small, would make it more likely than not *a posteriori* that β is positive. Such a prior, however, supposes that father's education is fairly close to being a valid instrument *and* that ability is only weakly correlated with schooling. If instead one took what we consider to be the more reasonable view that father's education is likely to be a very poor instrument, say $\rho_{uz} > 0.15$, the model and data would force one to conclude that the returns to schooling are necessarily negative!

We have found a clear red flag; the reasonable beliefs that $\kappa > 0.8$, $\rho_{T^*u} \in [0, 0.8]$, and $\rho_{uz} > 0$ are incompatible with the even more reasonable belief that $\beta > 0$. Moreover, the weaker beliefs that $\kappa > 0.8$ and $\rho_{T^*u} \in [0, 0.8]$ make $\beta > 0$ very unlikely for any reasonable prior over the identified set. As some have suggested that schooling and ability are *negatively* correlated (e.g. Ashenfelter and Krueger, 1994), one might wonder if imposing $\rho_{T^*u} < 0$ rather than the reverse would avoid this stark contradiction. Inspection of Figures 3a and 3b shows that such a belief would indeed be compatible with a positive β. Nevertheless, notice that this would require $\rho_{uz} < 0$, namely the very implausible belief that father's schooling is *negatively* correlated with son's ability. In this example, our framework is useful in pointing out that something is seriously amiss. Among the possibilities, heterogeneous effects may be important, measurement error may be far from classical, or the model may be seriously misspecified. In any case, the researcher may want to go back to the drawing board.

4.3 Was Weber Wrong?

In sharp contrast to the previous example, we now present an application where our framework leads to very different conclusions. Becker and Woessmann (2009) are interested in the long-run effect of the adoption of Protestantism in 16th Century Prussia on literacy rates. The paper exploits variation across counties in distance to Wittenberg, the city where Martin Luther introduced his ideas and preached, as an instrument for the Protestant share of the Population in the 1870s. Their main equation is

$$\text{Literacy rate} = \text{constant} + \beta\,(\text{Protestant share}) + \mathbf{x}'\gamma + u$$

and the first stage is

$$\text{Protestant Share} = \text{constant} + \pi\,(\text{Distance to Wittenberg}) + \mathbf{x}'\delta + v$$

Under this specification, Becker and Woessmann (2009) obtain an OLS estimate of $\hat{\beta}_{OLS} = 0.099$ (S.E = 0.010) and a twice-as-large IV estimate of $\widehat{\beta}_{IV} = 0.189$ (S.E. = 0.028). Throughout their main specifications, the authors include a set of covariates $\mathbf{x}$, which includes the fraction of the population younger than age 10, of Jews, of females, of individuals born in the municipality, of individuals of Prussian origin, the average household size, log population, population growth in the preceding decade, and the fraction of the population with unreported education information. In our exercise below we include this same set of controls as described in section 3.

Becker and Woessmann (2009) express beliefs about the three key parameters in our framework. First, their IV strategy relies on the assumption that $\rho_{uz} = 0$. Furthermore, despite using demographic data from the 1870s, the authors argue that the 1870 Prussian Census is regarded by historians to be highly accurate. As such, they express the belief that measurement error in literacy rates is small. Finally, Becker and Woessmann (2009) go through a lengthy discussion of the nature of the endogeneity of the Protestant share. Overall, they argue it is likely that Protestantism is *negatively* correlated with the error:

> ... wealthy regions may have been less likely to select into Protestantism at the time of the Reformation because they benefited more from the hierarchical Catholic structure, because the opportunities provided by indulgences allured to them , and because the indulgence costs weighted less heavily on them... The fact that "Protestantism" was initially a "protest" movement involving peasant uprisings that reflected social discontent is suggestive of such a negative selection bias. (pp. 556-557)

Notice that under the belief that measurement error is small, a negative correlation between the Protestant share and the error is necessary for the OLS estimate to be downward biased if the instrument is valid. A natural question in this example is, thus, whether the assumption that $\rho_{T^*u} < 0$ is compatible with little measurement error and a valid instrument. Figure 5a plots the full identified set for this exercise at the MLE for the data. One key feature of the identified set is that the constraint for measurement error from equation 19 binds. The data *alone* rules out a signal to noise ratio κ in the Protestant share below 0.5. This means that there is more signal than noise in measured Protestant share, but the data alone allows for the possibility of significantly more measurement error than the authors appear to entertain. The figure also depicts in green the subset of the identified set for which Becker and Woessmann (2009)'s prior that $\rho_{T*u} < 0$ holds.[15] Figure 5b also presents the unrestricted identified set at the MLE, this time depicting in blue the points for which the causal effect of Protestantism on literacy is positive ($\beta > 0$). A comparison of Figures 5a and 5b suggest what Figure 5c then clearly illustrates; under the author's prior and the restriction on κ implied by the data, the restricted identified set at the MLE rules out a non-positive β. Under a conditionally uniform prior over this restricted identified set, the left panel of Figures 6c and 6d present the corresponding densities for β, with and without accounting for sampling variability in Σ, which not only has a strictly positive support but also suggests the IV estimate may be *smaller* than the true causal effect. Note, however, that we have placed no restrictions on κ beyond those implied by the data. In particular, Figures 6c and 6d include values for β that correspond to κ as small as 0.5. This allows for the possibility of even greater measurement error than was present in the institutions variable from our first example, based on Acemoglu et al. (2001). This seems extreme given Becker and Woessmann's emphasis on the accuracy of their census data. Thus Figures 6e and 6f further restrict attention to those values of β that correspond to a stronger prior in which $\kappa > 0.8$. This allows at most 20% of the variation in measured Protestant share to be noise. Remarkably, the posterior for β under a conditionally uniform prior now aligns much more closely with the IV estimate and excludes nearly all of the implausibly large causal effects from Figures 6c and 6d. These very large values for β were only possible if we were willing to allow for the possibility of very severe measurement error in Protestant share.

It is also worth noting that, with or without a prior restriction on κ, the restricted identified set is fully compatible with a valid instrument and always rules out negative values for β. In this example, the authors beliefs are mutually consistent and their result is extremely robust, even to substantial instrument invalidity.

[15]More precisely we introduce the constraint that $-0.9 < \rho_{T*u} < 0$ to rule out outlandishly large negative correlations between Protestantism and the error

5 Conclusion and Extensions

In this paper we have presented a Bayesian procedure for inference in a non-identified linear model that is potentially subject to measurement error, an endogenous regressor, and an invalid instrument. Using three well-known examples from the empirical microeconomics literature, we have shown that our method can be highly informative, both about the mutual coherence of researcher beliefs and about the causal effect of interest.

We are currently working on two extensions to the results presented above. The first allows for multiple instrumental variables. While there are no serious technical obstacles to this extension, eliciting beliefs over multiple potentially invalid instruments may be difficult in practice. A second, more involved extension considers the case in which T^* is *binary*, a common situation in applied work. In this case, classical measurement error is impossible: if $T^* = 1$ then $w \in \{-1, 0\}$ while if $T^* = 0$ then $w \in \{0, 1\}$. This implies that $Cor(w, T^*) \leq 0$. Accordingly, in this case we assume that the measurement error is *non-differential* rather than classical: while the joint distribution of T^* and T is unrestricted, we assume that T is conditionally independent of all other variables in the system, given T^*.

When T^* is binary the problem becomes substantially more complicated, as the instrumental variables estimator does *not* eliminate the inconsistency arising from non-differential measurement error.[16] Moreover, in place of a single measurement error parameter κ, the binary case has two mis-classification probabilities $\alpha_0 = P(T = 1|T^* = 0)$ and $\alpha_1 = P(T = 0|T^* = 1)$. Adding these to ρ_{uz} and ρ_{T^*u} yields an identified set that comprises a three-dimensional manifold in four-dimensional space. Nevertheless, the same approach we outlined above can still be applied, although the details are somewhat different. Indeed, it may be even *easier* to elicit researcher beliefs over α_0, α_1 than κ. For example, in a regression of birth-weight on mother's smoking behavior, we would expect α_0, the fraction of non-smokers who claim to be smokers, to be nearly zero. Work on this and other empirical examples for the binary case is currently in progress.

[16] See for example Kane et al. (1999), Black et al. (2000) and Frazis and Lowenstein (2003).

References

Acemoglu, D., Johnson, S., Robinson, J. A., 2001. The colonial origins of comparative development: An empirical investigation. The American Economic Review 91 (5), 1369–1401.

Apostol, T. M., 1969. Calculus, 2nd Edition. Vol. II. John Wiley and Sons, New York.

Ashenfelter, O., Krueger, A., 1994. Estimates of the economic return to schooling from a new sample of twins. American Economic Review 84 (5), 1157–1173.

Becker, S. O., Woessmann, L., 2009. Was weber wrong? a human capital theory of protestant economic history. Quarterly Journal of Economics 124 (2), 531–596.

Berhman, J., Rosenzweig, M., 1999. "ability biases in schooling returns and twins: a test and new estimates. Economics of Education Review 18, 159–167.

Black, D. A., Berger, M. C., Scott, F. A., 2000. Bounding parameter estimates with nonclassical measurement error. Journal of the American Statistical Association 95 (451), 739–748.

Blackburn, M., Neumark, D., 1992. Unobserved ability, efficiency wages, and interindustry wage differentials. The Quarterly Journal of Economics 107 (4), 1421–1436.

Card, D., 1999. The causal effect of education on earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics. Vol. 3. Elsevier Science, Ch. 30.

Conley, T. G., Hansen, C. B., Rossi, P. E., 2012. Plausibly exogenous. The Review of Economics and Statistics 94 (1), 260–272.

Erickson, T., July 1993. Restricting regression slopes in the errrors-in-variables model by bounding the error correlation. Econometrica 61 (4), 959–969.

Frazis, H., Lowenstein, M. A., 2003. Estimating linear regressions with mismeasured, possibly endogenous, binary explanatory variables. Journal of Econometrics 117 (1), 151–178.

Griliches, Z., 1979. Sibling models and data in economics: Beginnings of a survey. Journal of Political Economy 87 (5), S37–S65.

Gustafson, P., May 2005. On model expansion, model contraction, identifiability and prior information: Two illustrative examples involving mismeasured variables. Statistical Science 20 (2), 111–140.

Gustafson, P., 2015. Bayesian Inference for Partially Identified Models: Exploring the Limits of Limited Data. No. 141 in Monographs on Statistics and Applied Probability. CRC Press, Boca Raton.

Hahn, P. R., Murray, J. S., Manolopoulou, I., March 2015. A bayesian partial identification approach to inferring the prevalence of accounting misconduct, working Paper.
URL http://arxiv.org/pdf/1407.8430

Kane, T., Rouse, C. E., Staiger, D., July 1999. Estimating the returns to schooling when schooling is misreported.

Kitagawa, T., July 2012. Estimation and inference for set-identified parameters using posterior lower probability, working Paper.
URL http://www.homepages.ucl.ac.uk/~uctptk0/Research/LowerUpper.pdf

Kline, B., Tamer, E., February 2015. Bayesian inference in a class of partially identified models, working Paper.
URL http://scholar.harvard.edu/files/tamer/files/kt-bayesian.pdf

Melfi, G., Schoier, G., 2004. Simulation of random distributions on surfaces. Societa Italiana di Statistica, 173–176.

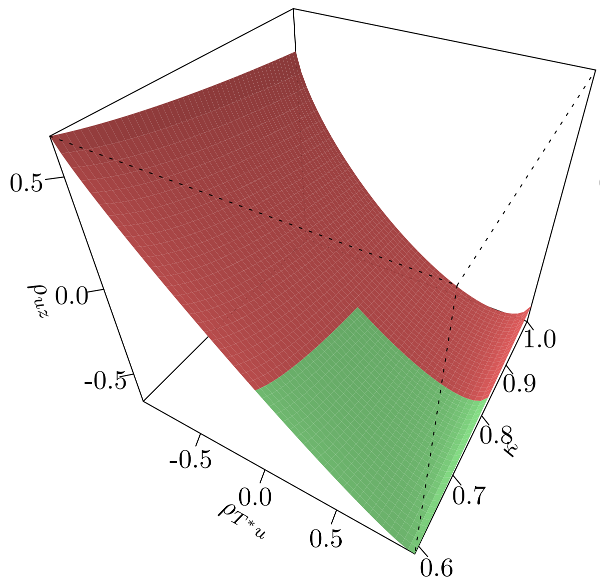
Moon, H. R., Schorfheide, F., 2009. Estimation with overidentifying inequality moment conditions. Journal of Econometrics 153, 136–154.

Moon, H. R., Schorfheide, F., 2012. Bayesian and frequentist inference in partially identified models. Econometrica 80 (2), 755–782.

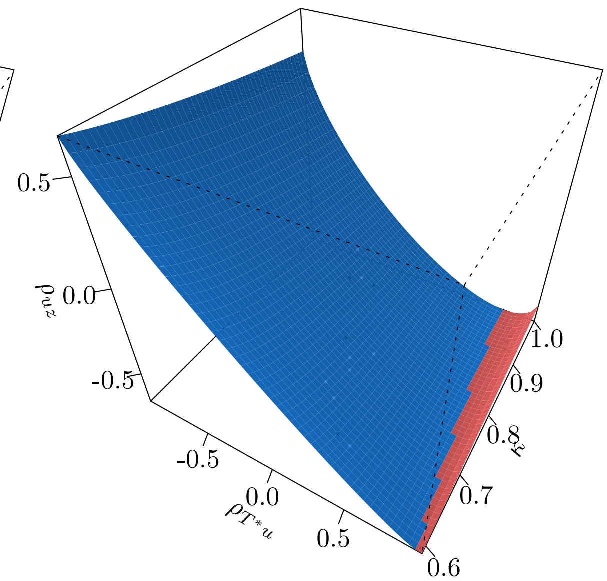
Nevo, A., Rosen, A. M., 2012. Identification with imperfect instruments. The Review of Economics and Statistics 94 (3), 659–671.

Poirier, D. J., 1998. Revising beliefs in nonidentified models. Econometric Theory 14, 483–509.

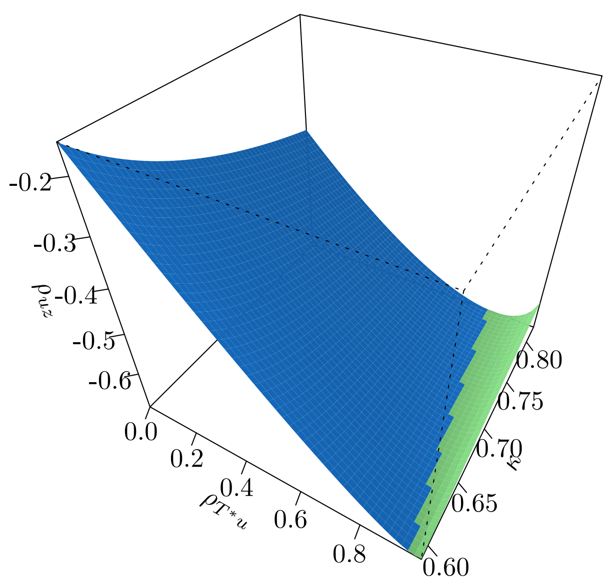
Richardson, T. S., Evans, R. J., Robins, J. A., 2011. Transparent parameterizations of models for potential outcomes. In: Bayesian Statistics. Vol. 9. pp. 569–610.



(a) Full, Prior in Green

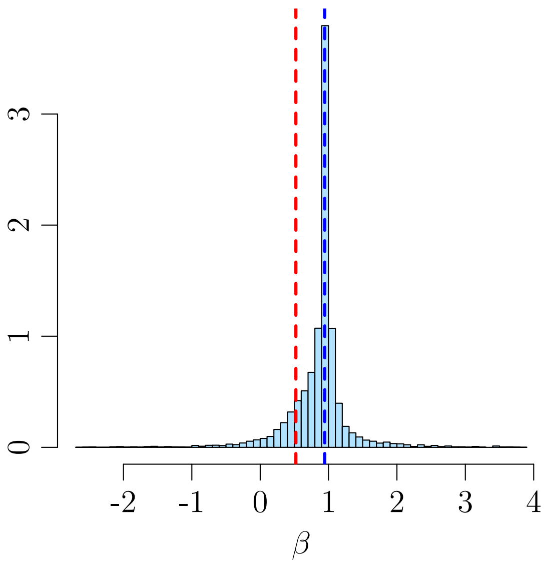


(b) Full, Positive β in Blue

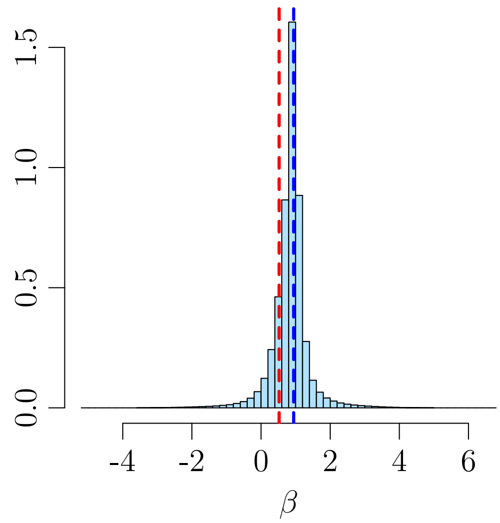


(c) Restricted, Positive β in Blue

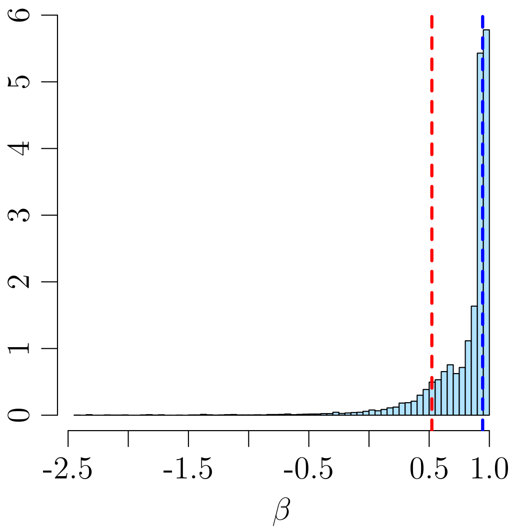
Figure 1: Identified set for the Colonial Origins example from Section 4.1. Panels (a) and (b) depict the full identified set for $(\rho_{T^*u}, \rho_{uz}, \kappa)$ evaluated at the MLE for Σ, while panel (c) restricts attention to the region $\kappa \in [0.45, 0.82], \rho_{T^*u} > 0$.



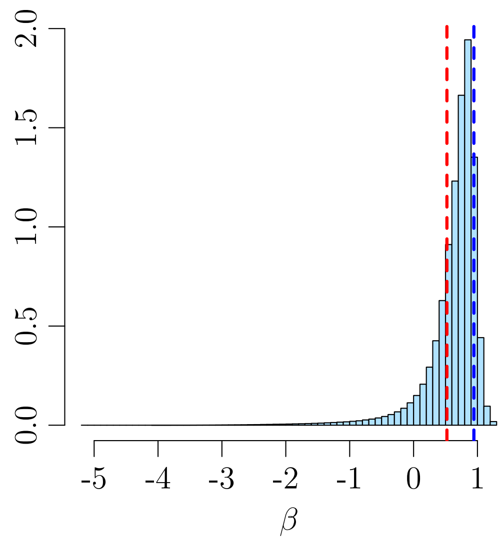
(a) Full, MLE



(b) Full, Posterior

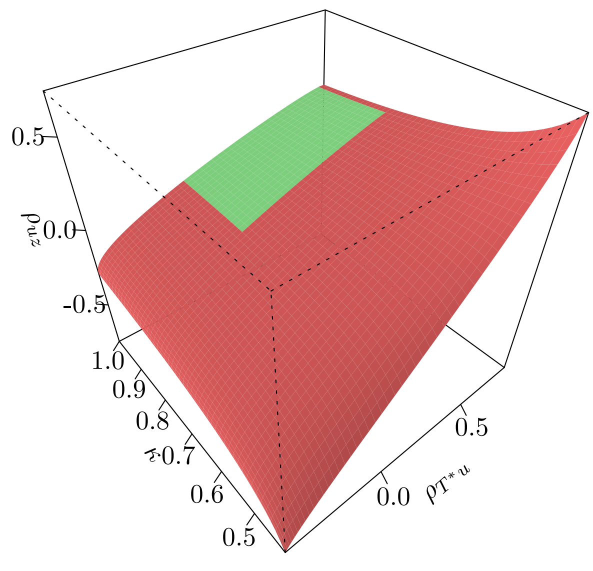


(c) Restricted, MLE

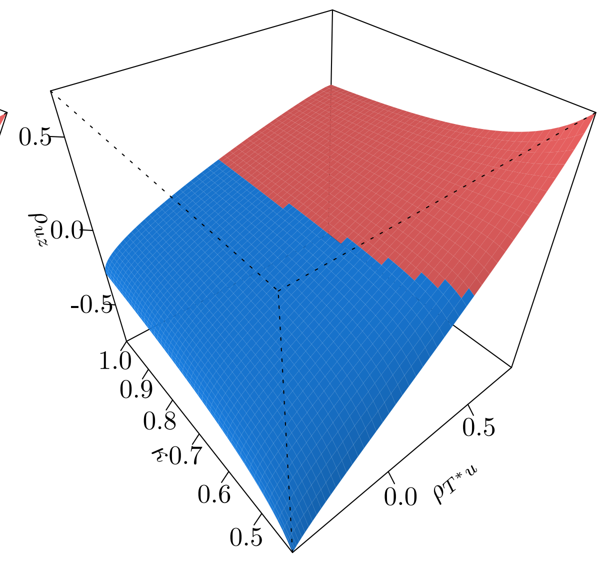


(d) Restricted, Posterior

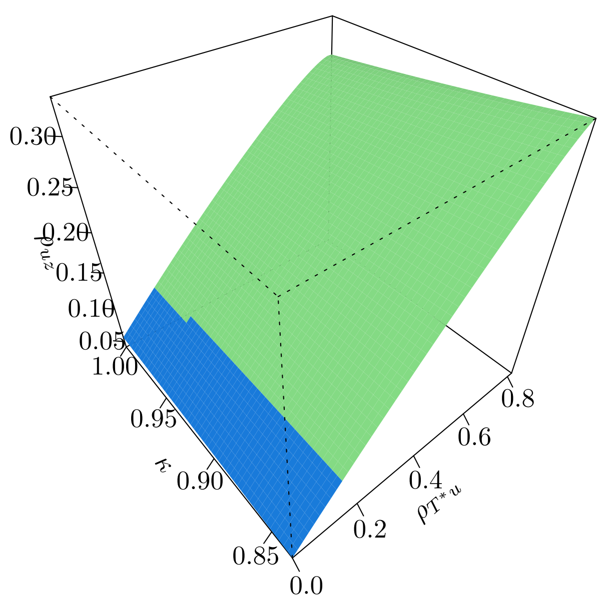
Figure 2: Histograms for β for the Colonial Origins example from Section 4.1. In each panel the dashed red line indicates the OLS estimate, and the dashed blue line the IV estimate. Panels (a) and (b) impose no prior restrictions on the identified set for $(\rho_{T^*u}, \rho_{uz}, \kappa)$, while panels (c) and (d) restrict attention to the region $\kappa \in [0.45.0.62], \rho_{T^*u} > 0$. Panels (a) and (c) are evaluated at the MLE for Σ, whereas panels (b) and (d) average over the posterior draws for $\Sigma^{(j)}$.



(a) Full, Prior in Green

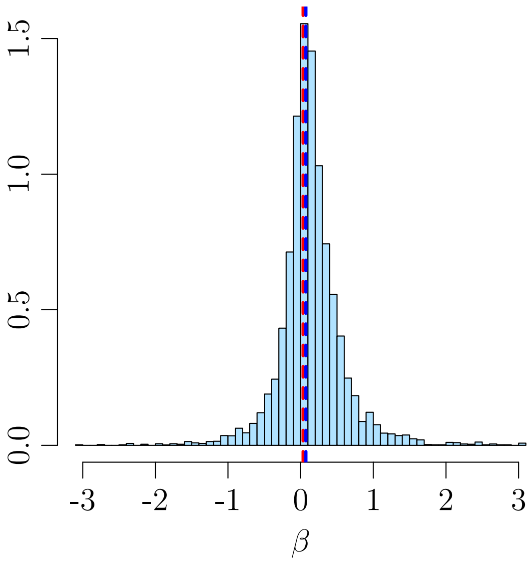


(b) Full, Positive β in Blue

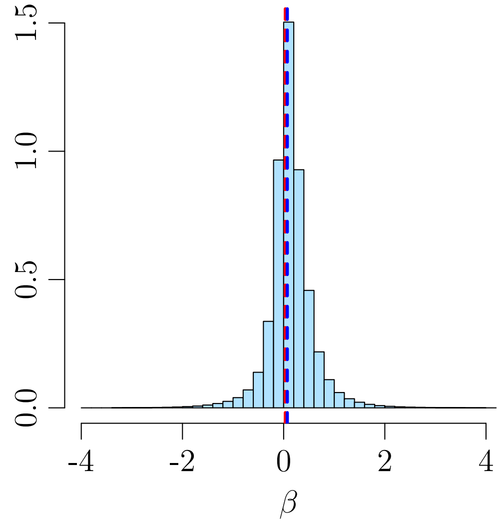


(c) Restricted, Positive β in Blue

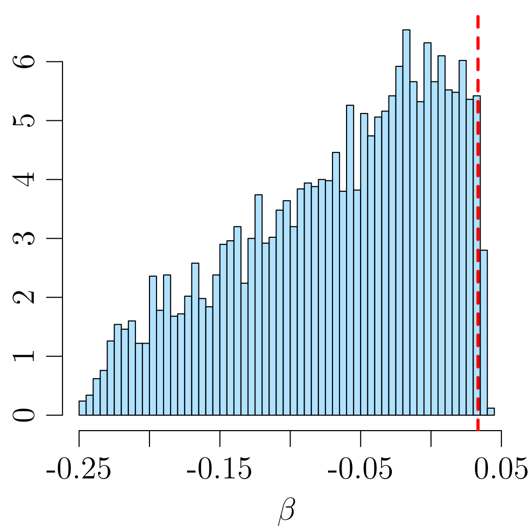
Figure 3: Identified set for Returns to Schooling example from Section 4.2. Panels (a) and (b) depict the full identified set for $(\rho_{T^*u}, \rho_{uz}, \kappa)$ evaluated at the MLE for Σ, while panel (c) restricts attention to the region that satisfies the prior $\kappa > 0.8, \rho_{T^*u} \in [0, 0.8]$.



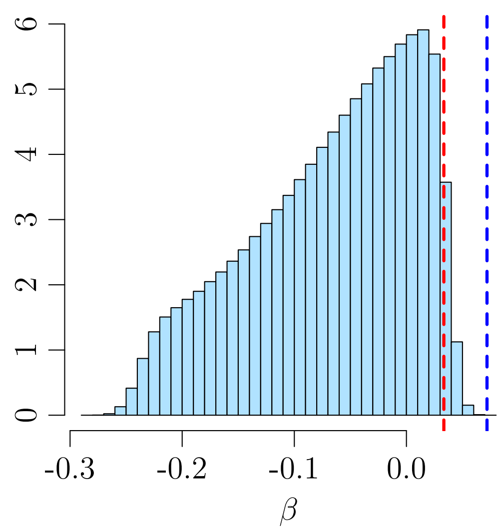
(a) Full, MLE



(b) Full, Posterior

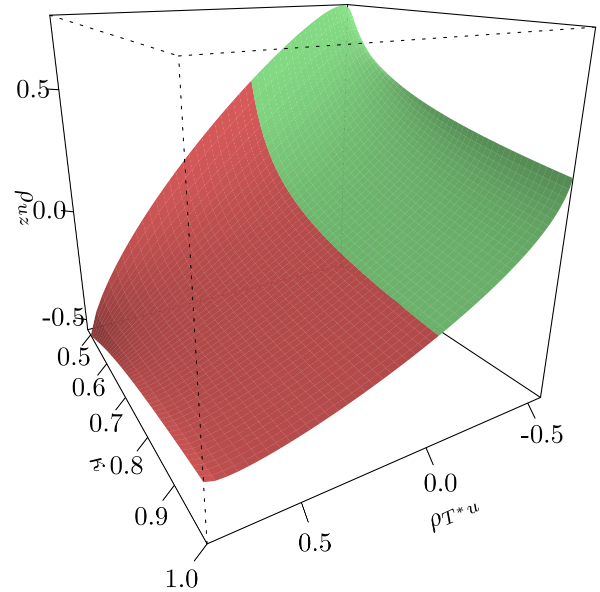


(c) Restricted, MLE

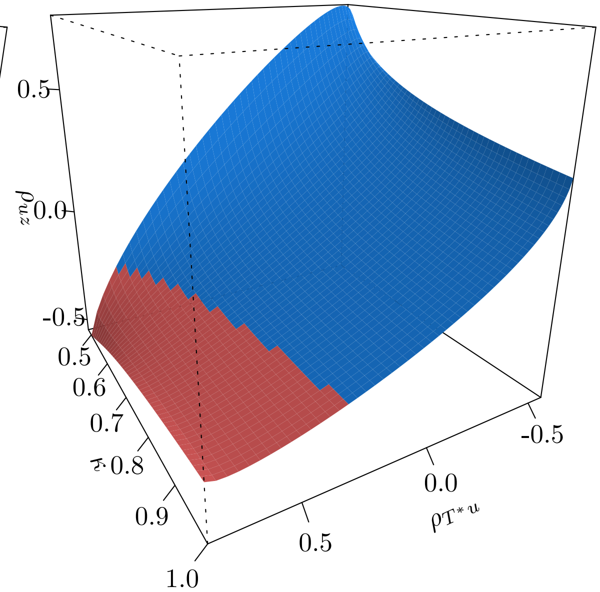


(d) Restricted, Posterior

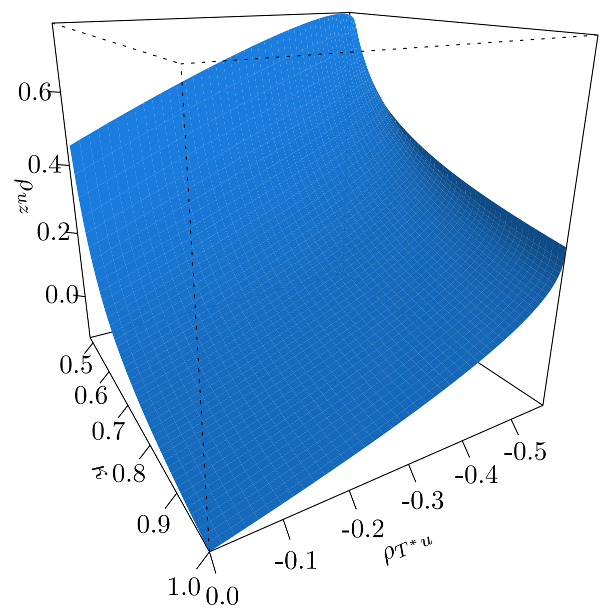
Figure 4: Histograms for β for the Returns to Schooling example from Section 4.2. In each panel the dashed red line indicates the OLS estimate, and the dashed blue line the IV estimate. Panels (a) and (b) impose no prior restrictions on the identified set for $(\rho_{T^*u}, \rho_{uz}, \kappa)$, while panels (c) and (d) restrict attention to the region that satisfies the prior $\kappa > 0.8, \rho_{T^*u} \in [0, 0.8]$. Panels (a) and (c) are evaluated at the MLE for Σ, whereas panels (b) and (d) average over the posterior draws for $\Sigma^{(j)}$.



(a) Full, Prior in Green

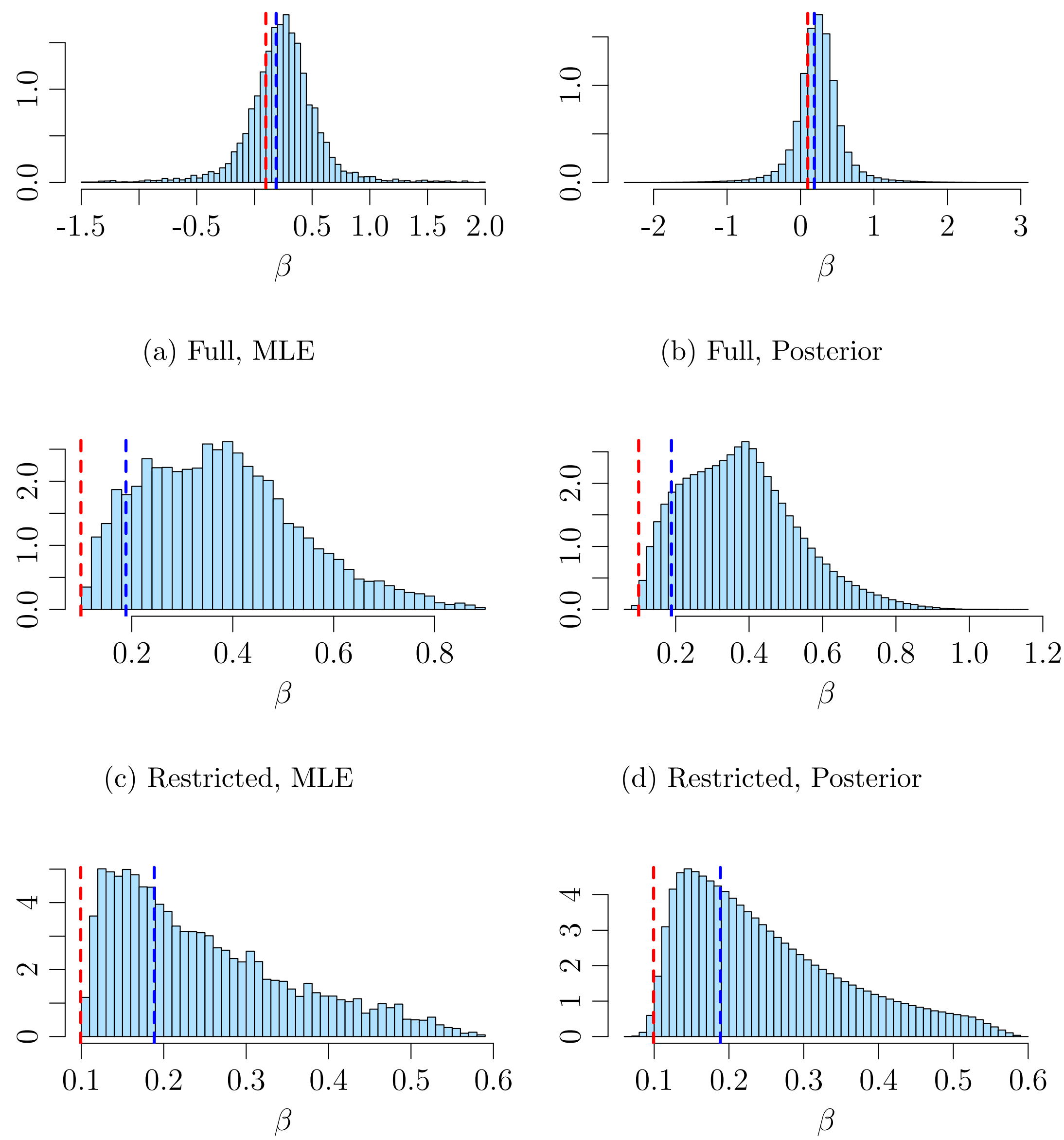


(b) Full, Positive β in Blue



(c) Restricted, Positive β in Blue

Figure 5: Identified set for the "Was Weber Wrong?" example from Section 4.3. Panels (a) and (b) depict the full identified set for $(\rho_{T^*u}, \rho_{uz}, \kappa)$ evaluated at the MLE for Σ, while panel (c) restricts attention to the region that satisfies the prior $\rho_{T^*u} < 0$.



(a) Full, MLE

(b) Full, Posterior

(c) Restricted, MLE

(d) Restricted, Posterior

(e) Restricted, MLE

(f) Restricted, Posterior

Figure 6: Histograms for β for the "Was Weber Wrong?" example from Section 4.3. In each panel the dashed red line indicates the OLS estimate, and the dashed blue line the IV estimate. Panels (a) and (b) impose no prior restrictions on the identified set for $(\rho_{T^*u}, \rho_{uz}, \kappa)$. Panels (c) and (d) restrict attention to the region that satisfies the prior $\rho_{T^*u} < 0$ while Panels (e) and (f) add the further restriction that $\kappa > 0.8$. Panels (a), (c) and (e) are evaluated at the MLE for Σ, whereas panels (b), (d) and (f) average over the posterior draws for $\Sigma^{(j)}$.