# Is the Great Gatsby Curve Robust?

## Comment on Corak (2013)

Bradley J. Setzler[1]

Presented to Economics 350
Department of Economics
University of Chicago
setzler@uchicago.edu

January 15, 2014

[1]Thanks to James Heckman for many helpful comments.

# Introduction

Intergenerational Elasticity of Earnings: The coefficient $\beta$ in the regression,

$$\ln Y_{i,j}^{c} = \alpha_j + \beta_j \ln Y_{i,j}^{p} + \delta_j X_{i,j} + \epsilon_{i,j}, \tag{1}$$

where $Y^p$ is parents' lifetime earnings, $Y^c$ is children's lifetime earnings, $X$ is a vector of covariates often including a polynomial in parent's age ($T$), $j$ is the index over countries, and $i$ is the index over individuals. See Black and Devereaux (2011) for a recent survey

# Introduction

The Great Gatsby Curve: the relationship between inequality and the intergenerational elasticity of earnings (IGE).
Popularized in a speech by Alan Krueger (2012), it is defined by Corak (2013) as the regression,

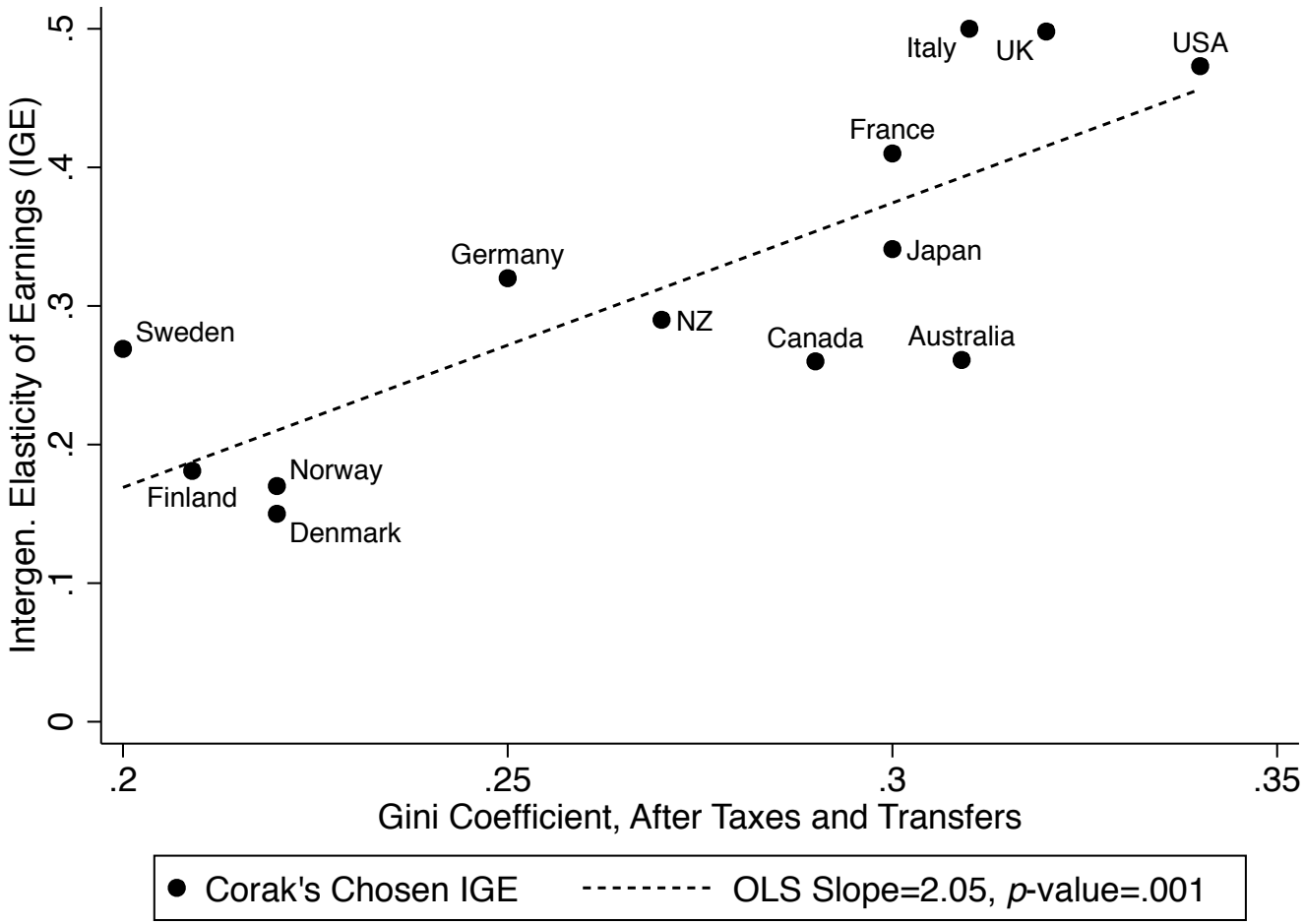
$$\beta_j = \gamma_0 + \gamma_1 G_j + \eta_j, \tag{2}$$

where $G$ is the disposable income Gini coefficient (a measure of inequality) and $\beta$ is estimated from Equation (1).[2]

[2] No explicit reference to the error $\eta$ is made in the literature, but this error structure is presupposed by the OLS estimation employed. Left implicit is the least strong assumption needed for unbiasedness and consistency of OLS, that $G$ and $\eta$ are uncorrelated. In an open economy, I expect this conditon to be violated as omitted variables (e.g., labor supply, the skill distribution, technology) are expected to causally affect both inequality and earnings, and to vary across countries. Nevertheless, I suspend disbelief in this comment.

# Introduction

Figure : The Great Gatsby Curve of Corak (2013)



# Introduction

Corak finds that, as a country becomes more unequal, earnings become more strongly preserved from parents to children. This result would raise serious questions about intergenerational distributive justice.

But where did the data come from? Are the subjective aspects of data selection driving the inference that the Great Gatsby Curve has positive slope? I test the sensitivity of the inference to Corak's choices.

> *A fragile inference is not worth taking seriously.*
> *-Ed Leamer (1985, AER)*

# Robustness Check 1: Gini Coefficients

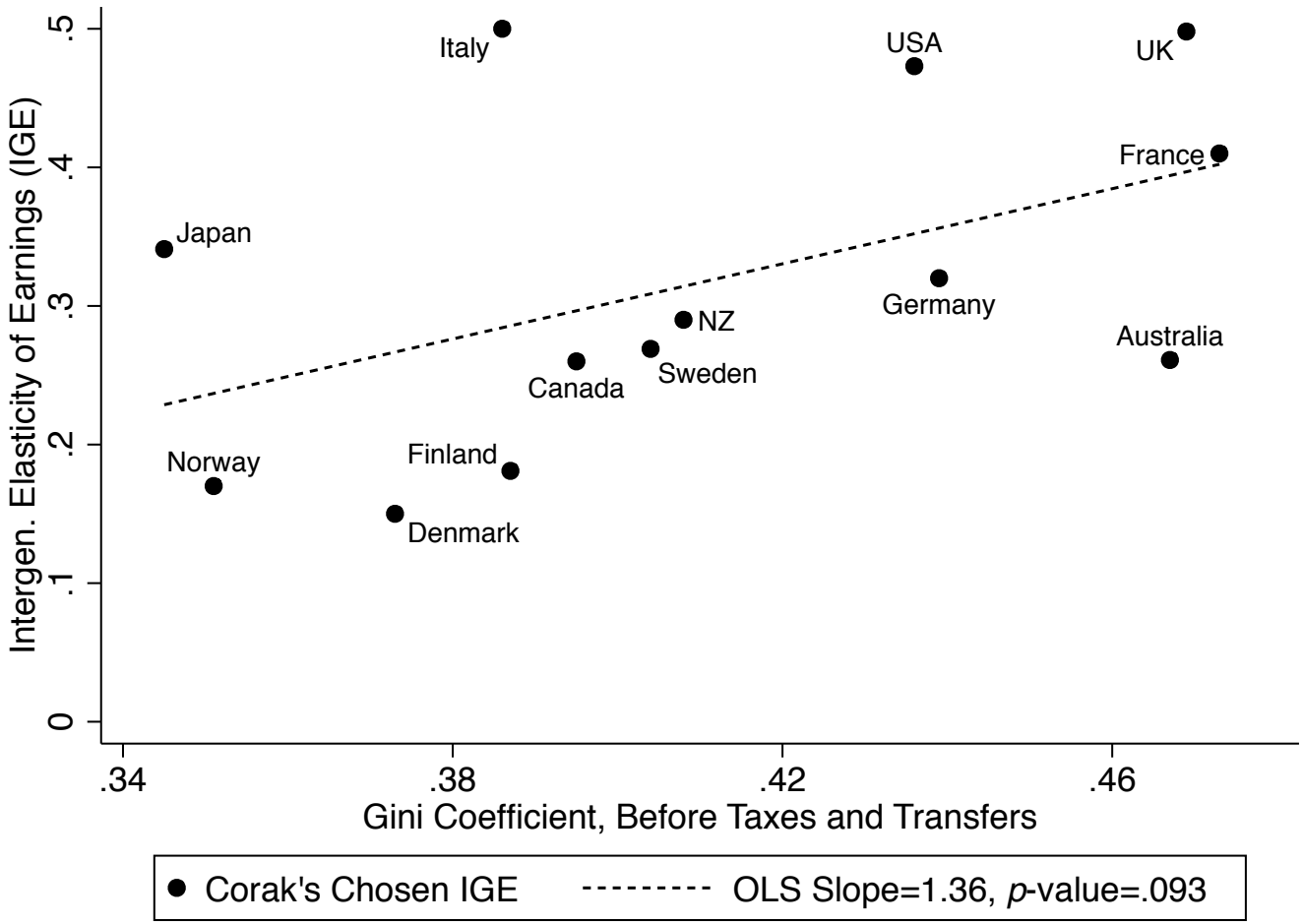
Corak fits the Great Gatsby Curve by regressing *before* tax-and-transfer earnings on *after* tax-and-transfer Gini coefficients. Recall Footnote 2: the model specification assumes that the IGE is linearly related to inequality, but it is difficult to imagine that,

1. taxes and transfers to do not affect earnings inequality (either in the same period or over time),
2. the relationship between taxes and transfers and earnings inequality is captured by subsuming taxes and transfers into the Gini coefficient only.

I compute the Great Gatsby Curve when *before* tax-and-transfer Gini coefficients are used. The result is striking: the Great Gatsby Curve no longer has statistically significant slope.

# Robustness Check 1: Gini Coefficients

Figure : Corak's Great Gatsby Curve (Before Tax-and-Transfer Gini)



# Robustness Check 2: Corak's Weights

Although Corak is collecting IGE estimates from other sources, he is not using them as reported. He uses the following process to modify the data prior to use:

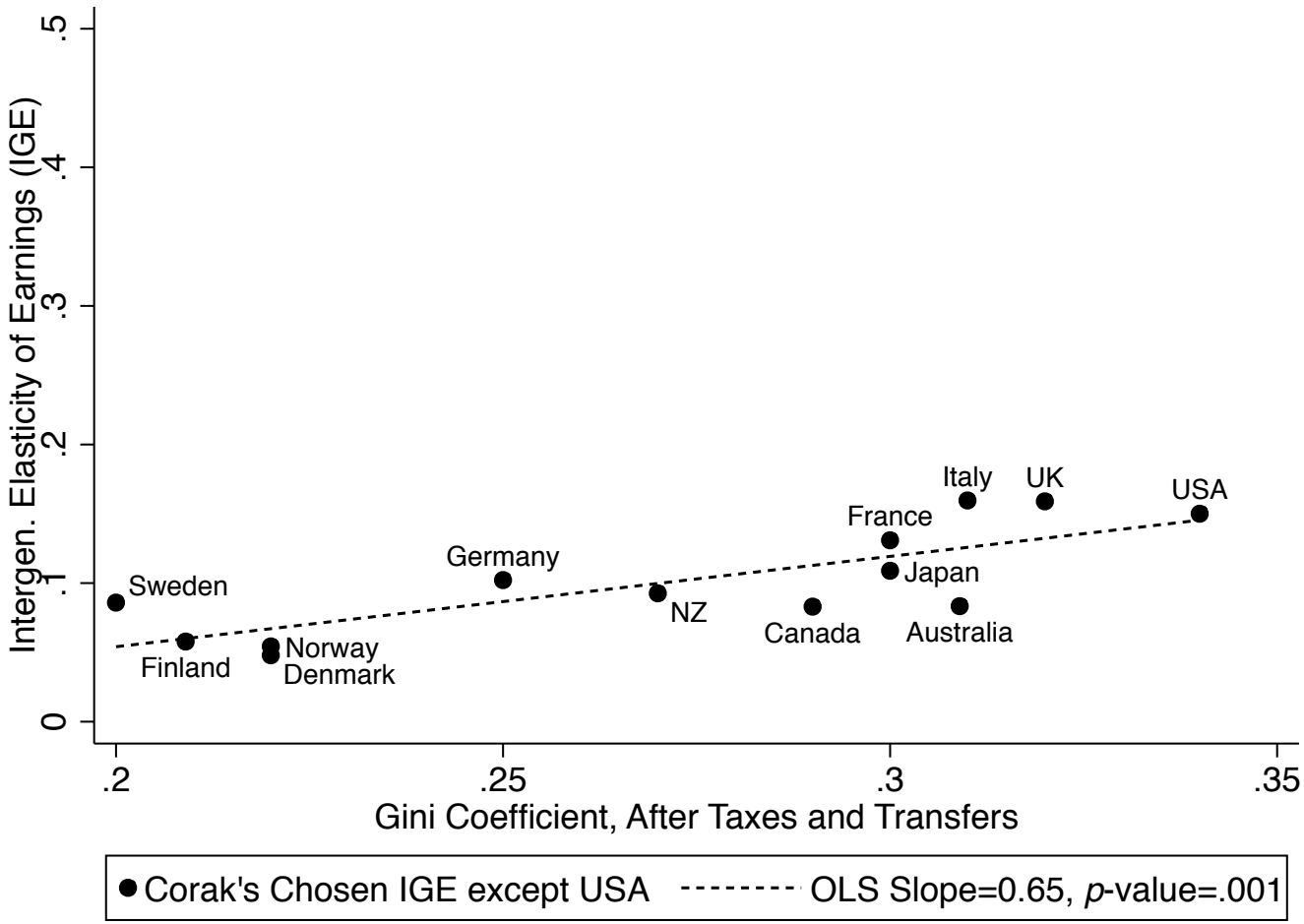
1. Pick the "most plausible" IGE estimate for the USA from among 41 published estimates.
2. Match every non-USA country to a USA IGE estimate from the literature based upon similarity of the estimation method.
3. If the matched USA IGE estimate is different from the "most plausible" estimate by a factor of $F$, then the corresponding non-USA IGE is modified by a factor of $F$.

# Robustness Check 2: Corak's Weights

Suppose that we disagree with Corak's (at least partly) subjective choice of the "most plausible" IGE estimate for the USA? Corak took his "most plausible" estimate from Grawe (2004). Grawe (2004) reported two values as *equally plausible*: Corak's chosen 0.47 as well as 0.15. I compute the Great Gatsby Curve using Grawe's lower estimate. However, I find that the Great Gatsby Curve is unaffected if we discard Corak's weights.

# Robustness Check 2: Corak's Weights

Figure : Corak's Great Gatsby Curve with Lower USA IGE Estimate



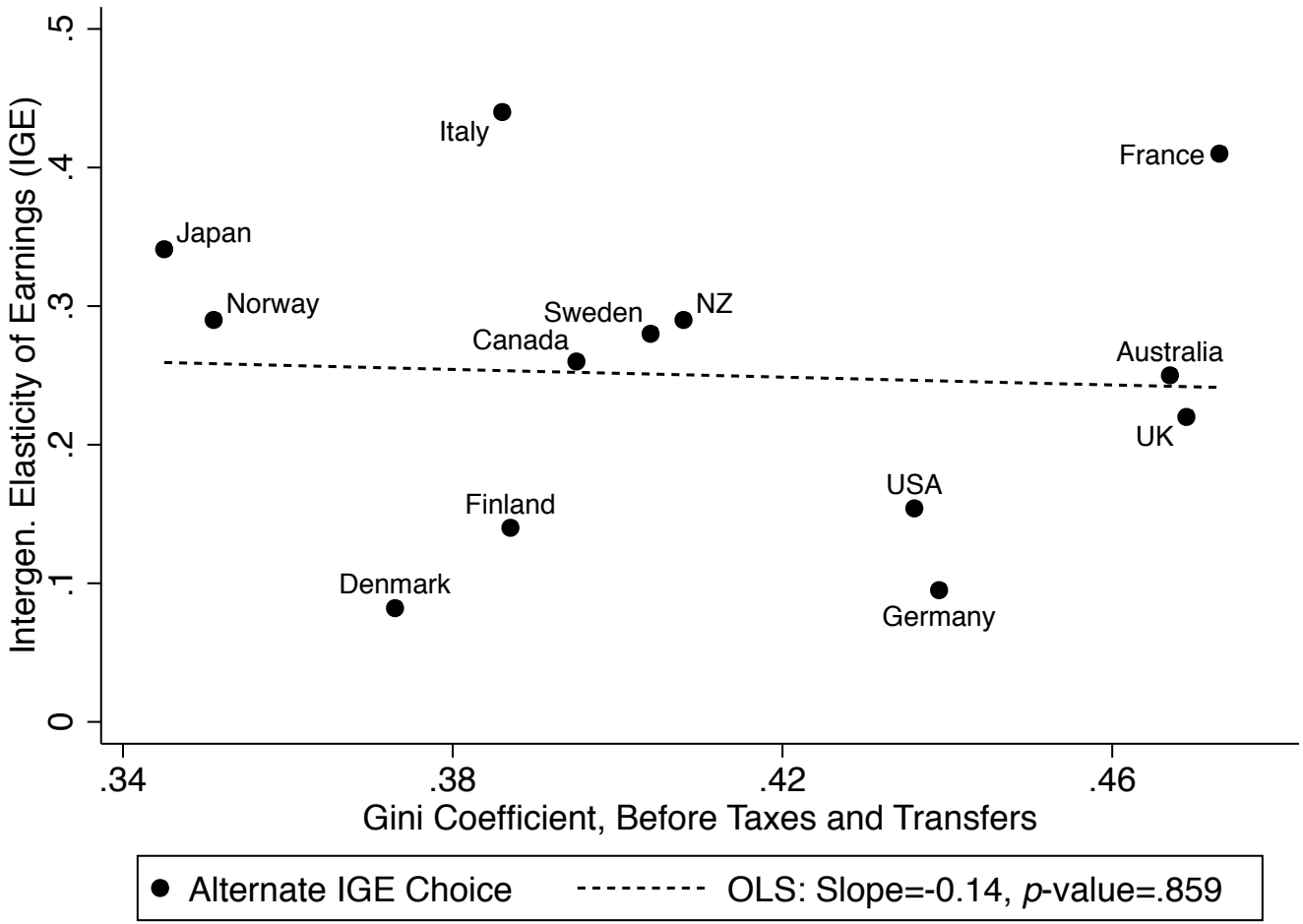
# Robustness Check 3: Corak's Chosen Data

We have seen how Corak's Great Gatsby Curve is affected if we maintain his weighting scheme but select an alternate USA IGE estimate: the slope shifts proportional to the change in USA IGE. We have also noted that we can discard his weighting scheme without qualitatively changing his results.

But how will the results change if we choose from among the many other IGE estimates available in the literature for the various countries? To motivate the question, I replace Corak's IGE choices for three of the 13 countries (Norway, UK, and USA). Two of my estimates are taken from papers from which Corak takes some of his chosen estimates. I find that the Great Gatsby Curve now has a *negative* slope.

# Robustness Check 3: Corak's Chosen Data

Figure : Great Gatsby Curve using Three Alternate Elasticities (Gini Before)



# Robustness Check 3: Corak's Chosen Data

More generally, consider the following sensitivity analysis: using a survey of every IGE estimate available on each country, estimate the Great Gatsby Curve for every combination. There are 246,000 combinations, so I obtain a distribution of 246,000 slope coefficient estimates and 246,000 corresponding $p$-values.

Using before tax-and-transfer Gini coefficients, 97.6% have positive estimated slope coefficient but only 0.003% of these slope estimates are statistically significant. The mean coefficient is 0.916 and the standard deviation is 0.467, so zero is just within two standard deviations of the mean.

Using after tax-and-transfer Gini coefficients, none has negative estimated slope coefficient and 83% of these slope estimates are statistically significant. The mean coefficient is 1.846 and the standard deviation is 0.509.

# Robustness Check 3: Corak's Chosen Data

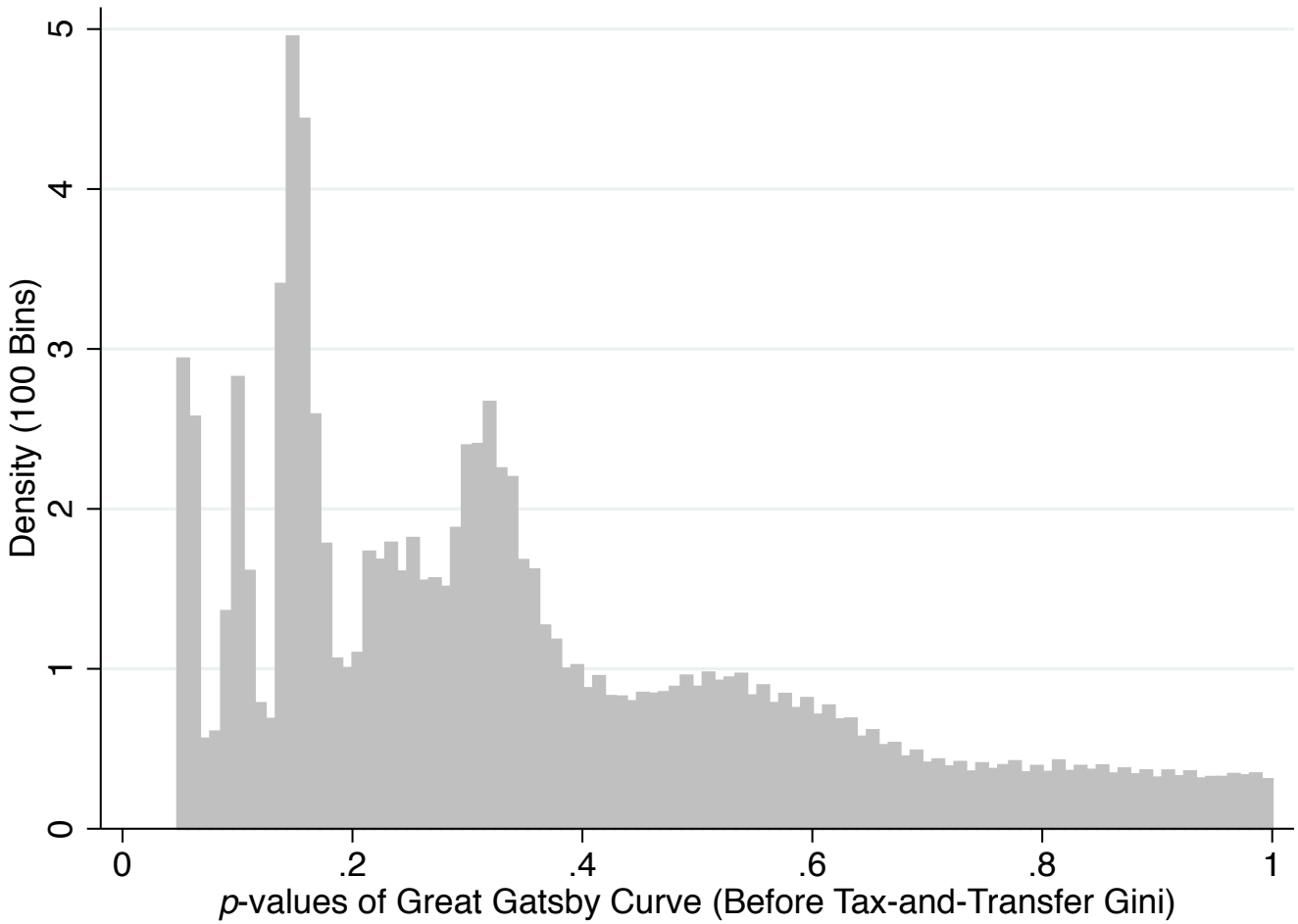
**My sensitivity analysis reveals that Corak has chosen IGE's that correspond to a Great Gatsby Curve slope estimate that is in the largest 15% using before tax-and-transfer Gini coefficients and in the largest 10% using after tax-and-transfer Gini coefficients.**

# Robustness Check 3: Corak's Chosen Data

Figure : Distribution of Great Gatsby Curve Slopes for All IGE Combinations

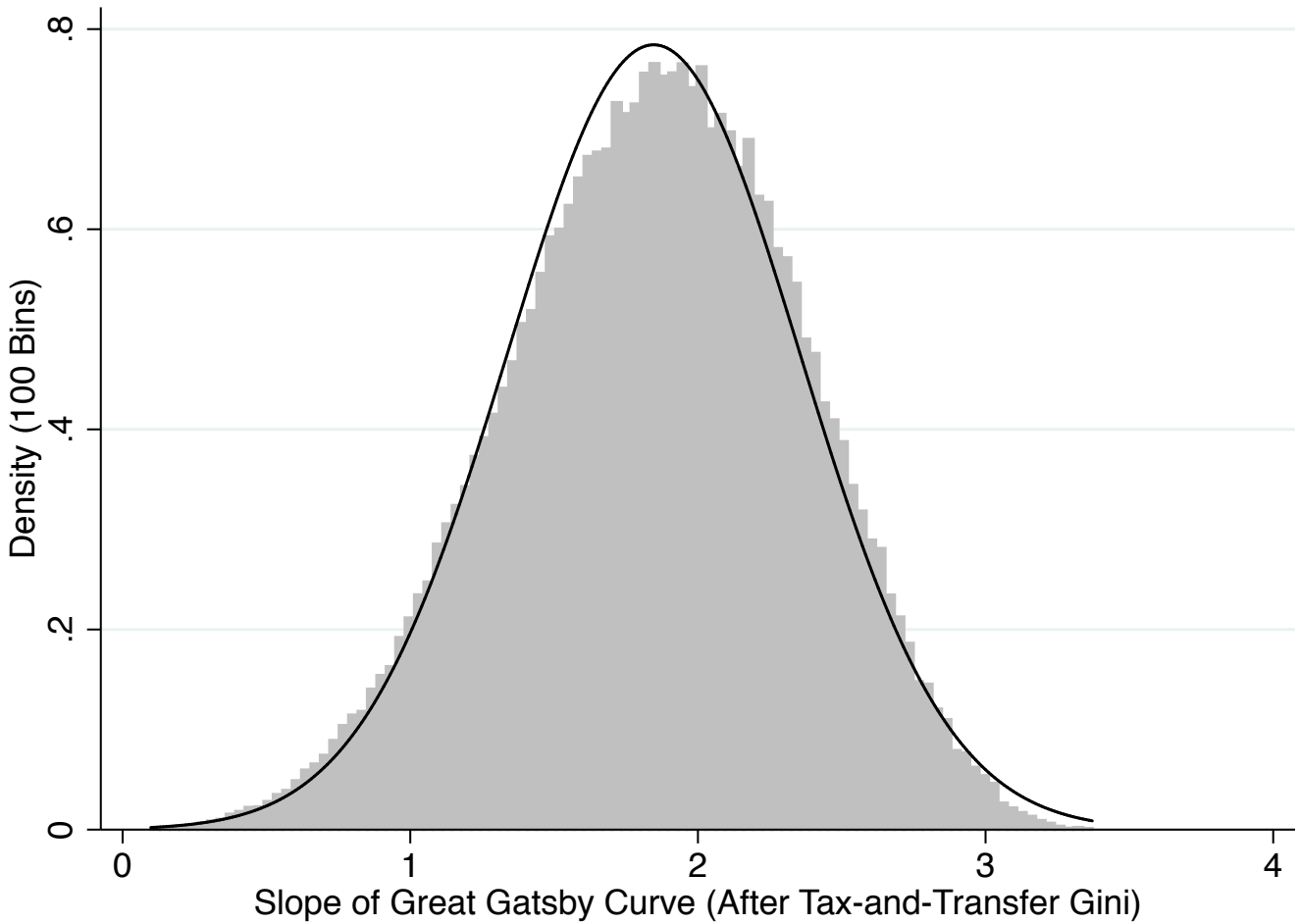
# Robustness Check 3: Corak's Chosen Data

Figure : Distribution of Great Gatsby Curve *p*-values for All IGE Combinations



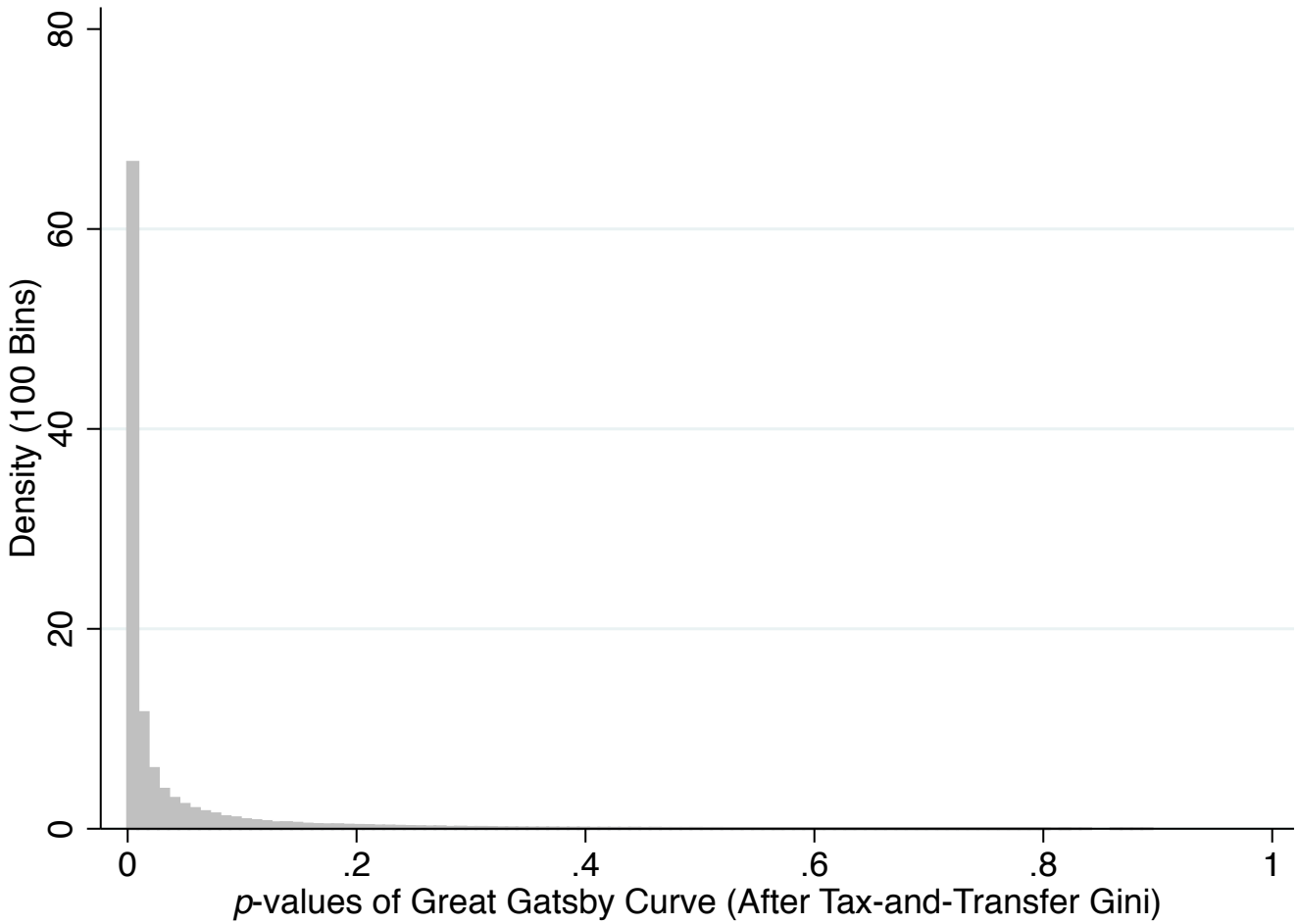
# Robustness Check 3: Corak's Chosen Data

Figure : Distribution of Great Gatsby Curve Slopes for All IGE Combinations



# Robustness Check 3: Corak's Chosen Data

Figure : Distribution of Great Gatsby Curve *p*-values for All IGE Combinations



# Conclusions

I have demonstrated that the leading estimate of the Great Gatsby Curve, the least-squares regression of Corak (2013), fails three tests of robustness:

1. It is not robust to the choice of Gini coefficient (before tax-and-transfer versus after tax-and-transfer.)
2. It is not robust in level of slope to the choice of USA IGE.
3. Even the sign of the slope is not robust to the choice of alternate IGE choices across countries when using before tax-and-transfer Gini coefficients.

I propose disentangling the effects of earnings inequality, taxes, and transfers on the IGE.[3]

[3] I continue to suspend disbelief regarding omitted variables and linear functional form, as described in Footnote 2.

# Conclusions

Using estimates by the OECD (2011) of average tax rate ($T$) payed by the top 20% of the earnings distribution and average benefit as a fraction of income ($B$) transferred to the bottom 20% of the earnings distribution, I fit all possible combinations of IGE estimates to the modified Great Gatsby Curve,[4]

$$\beta_j = \gamma_0 + \gamma_1 G_j^B + \gamma_2 T_j + \gamma_3 B_j + \eta_j, \tag{3}$$

where $G^B$ is the before tax-and-transfer Gini coefficient.

[4] Even more so than the Gini coefficients from the OECD (2011,2013), the tax and transfer estimates correspond to more recent years than the IGE estimates, so the validity of the estimates depends on similarity over time in the Gini coefficients and tax and transfer rates.

# Conclusions

I find that the mean coefficient is 1.428 and 37.7% are statistically significant. Furthermore, I find that the mean coefficients on taxes and transfers are -0.833 and -0.204 and 70.5% and 95.7% are statistically significant, respectively.

Thus, I conclude that inequality has a positive and somewhat statistically significant effect on IGE, while taxes and transfers have negative and mostly statistically significant effects on IGE. I have no reason to believe that this result is robust to alternate model or statistical assumptions. Take this only as motivation for a more general analysis of the relationship between inequality and intergenerational mobility across countries.

See my paper for more details and citations.