

## **EMPIRICAL DATA GATHERING ADDENDUM**

*A Practical Guide for Implementing the Empirical Strategy*

Companion to: "Taxing Concentration, Not Transfer"

### **1. OVERVIEW**

This addendum provides step-by-step instructions for gathering the empirical data needed to implement the estimation strategy in the main paper. The goal is to move from calibrated projections to actual empirical estimates.

The empirical strategy requires data for three purposes:

1. Structural estimation of behavioral parameters (bequest elasticity, planning effectiveness, dispersion preference)
2. Cross-country comparison (UK vs. US inheritance tax systems)
3. Within-US analysis (state inheritance tax variation)

#### **Estimated Timeline:**

- Public data assembly: 2-4 weeks
- Restricted data applications: 3-6 months for approval
- Data cleaning and preparation: 4-8 weeks
- Estimation and analysis: 8-12 weeks
- Total: 6-12 months for full implementation

### **2. US DATA SOURCES**

#### **2.1 IRS Statistics of Income (SOI) — Estate Tax Data**

##### **What it contains:**

Estate tax return data including gross estate value, deductions, taxable estate, tax liability, asset composition, and beneficiary information. Available at varying levels of detail.

##### **Public-Use Data (Free, Immediate Access):**

URL: <https://www.irs.gov/statistics/soi-tax-stats-estate-tax-statistics>

Contains: Aggregate statistics by estate size, year, state. Published tables going back to 1916.

Limitations: No microdata, no beneficiary-level detail.

##### **SOI Public-Use File (PUF):**

URL: <https://www.irs.gov/statistics/soi-tax-stats-estate-tax-study-public-use-file>

Contains: Individual estate-level data with identifying information removed.  
Available for selected years.

Variables: Gross estate, net estate, deductions by type, tax before/after credits, marital status, state.

Access: Free download after registration.

**Restricted Microdata (Research Access Required):**

Application: Submit proposal to IRS Statistics of Income Division

URL: <https://www.irs.gov/statistics/soi-tax-stats-statistics-of-income-research-program>

Contains: Full microdata including beneficiary information, trust details, specific asset values.

Requirements: Must be affiliated with research institution, proposal review takes 3-6 months.

Alternative: Access through Federal Statistical Research Data Centers (FSRDCs).

**Key Variables to Extract:**

Variable	SOI Field	Use in Analysis
Gross estate	GRSS_EST	Estate size distribution
Net estate tax	NET_EST_TX	Effective tax rate
Charitable deduction	CHAR_DED	Charitable response
Marital deduction	MARITAL_DED	Spousal transfers
State	STATE	State-level analysis
Year of death	YR_DEATH	Time series analysis

**2.2 Survey of Consumer Finances (SCF)**

**What it contains:**

Triennial survey of household wealth, income, and financial behavior. Includes questions on inheritances received and expected.

**Access:**

URL: <https://www.federalreserve.gov/econres/scfindex.htm>

Public-use data freely downloadable in Stata, SAS, and ASCII formats.

**Key Variables:**

- X5804-X5809: Inheritances received (amount, year, from whom)
- X5821-X5826: Expected inheritances
- NETWORTH: Total household net worth

- INCOME: Total household income
- Various demographic and financial variables

**Sample Code (Stata):**

```
use scf2022.dta, clear
* Identify households that received inheritance
gen received_inheritance = (x5804 > 0)
* Amount of largest inheritance
gen inheritance_amount = x5804
* Tabulate by wealth decile
xtile wealth_decile = networth, nq(10)
tab wealth_decile received_inheritance
```

**2.3 Panel Study of Income Dynamics (PSID)**

**What it contains:**

Longitudinal survey following families since 1968. Wealth supplements (1984, 1989, 1994, 1999, 2001, 2003, and biennially since) include inheritance questions.

**Access:**

URL: <https://psidonline.isr.umich.edu/>

Free registration required. Custom data extracts available through online tool.

**Advantage over SCF:**

Panel structure allows tracking wealth before and after inheritance receipt—essential for estimating causal effects on labor supply and consumption.

**Key Variables:**

- ER71485 (2019): Received inheritance/gift since last interview
- ER71486: Amount of inheritance
- ER71487: Year received
- ER71488: Relationship to person who left inheritance

**2.4 Forbes 400 Data**

**What it contains:**

Annual list of 400 wealthiest Americans with wealth estimates, source of wealth, and biographical information.

**Access:**

Historical data: <https://www.forbes.com/forbes-400/>

Academic compilations: Kaplan and Rauh (2013) provide cleaned panel data.

Peterson Institute database: <https://www.piie.com/publications/working-papers/origins-superrich-billionaire-characteristics-database>

**Use in Analysis:**

Track persistence of dynastic wealth at the very top. Identify inherited vs. self-made fortunes. Calculate implied bequest patterns.

### 3. STATE-LEVEL INHERITANCE TAX DATA

#### 3.1 Current State Inheritance Taxes

Six states currently levy inheritance taxes with rates varying by beneficiary class:

State	Spouse Rate	Child Rate	Sibling Rate	Other Rate
Iowa	Exempt	Exempt	5-10%	10-15%
Kentucky	Exempt	Exempt	4-16%	6-16%
Maryland	Exempt	Exempt	10%	10%
Nebraska	Exempt	1%	11%	15%
New Jersey	Exempt	Exempt	11-16%	15-16%
Pennsylvania	Exempt	4.5%	12%	15%

#### 3.2 Historical State Data

Before EGTRRA (2001) phased out the federal credit for state death taxes, most states had inheritance or estate taxes. Compiling historical tax parameters is essential for the diff-in-diff analysis.

**Primary Source:**

Commerce Clearing House (CCH) State Tax Handbook (annual editions)

Available in law libraries and through Wolters Kluwer subscription.

**Secondary Sources:**

- McIntyre, Michael J. "State Death Taxes After EGTRRA: A Long Day's Journey into Night." *State Tax Notes* (2002).
- Conway, Karen Smith, and Jonathan Rork. "State 'Death' Taxes and Elderly Migration." *National Tax Journal* (2004).
- NCSL State Tax Actions Database: <https://www.ncsl.org/research/fiscal-policy/state-tax-actions-database.aspx>

**Data to Compile:**

For each state-year (1990-2025):

- Type of tax (inheritance, estate, both, neither)

- Tax rates by beneficiary class
- Exemption amounts by beneficiary class
- Special provisions (family business exemptions, etc.)

### 3.3 State Revenue Data

#### **Census Bureau State Government Finances:**

URL: <https://www.census.gov/programs-surveys/state/data/tables.html>

Contains: Annual state tax revenue by type, including "death and gift taxes."

Years: 1992-present in consistent format.

#### **Individual State Revenue Departments:**

Many states publish detailed revenue statistics. Contact information:

- Iowa: [tax.iowa.gov](http://tax.iowa.gov)
- Kentucky: [revenue.ky.gov](http://revenue.ky.gov)
- Maryland: [marylandtaxes.gov](http://marylandtaxes.gov)
- Nebraska: [revenue.nebraska.gov](http://revenue.nebraska.gov)
- New Jersey: [nj.gov/treasury](http://nj.gov/treasury)
- Pennsylvania: [revenue.pa.gov](http://revenue.pa.gov)

## 4. UK DATA FOR CROSS-COUNTRY ANALYSIS

### 4.1 HMRC Inheritance Tax Statistics

#### **What it contains:**

Annual statistics on inheritance tax including estates chargeable, tax paid, reliefs claimed, and distributions by estate size.

#### **Access:**

URL: <https://www.gov.uk/government/statistics/inheritance-tax-statistics>

Free download of published tables and commentary.

#### **Key Tables:**

- Table 12.1: Estates notified and tax due by estate size
- Table 12.2: Estates by type of asset
- Table 12.3: Reliefs and exemptions claimed
- Table 12.4: Time series of IHT statistics

### 4.2 Wealth and Assets Survey (WAS)

#### **What it contains:**

Biennial survey of household wealth in Great Britain, analogous to US SCF.  
Includes inheritance questions.

**Access:**

URL: <https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/debt/methodology>

Public-use data through UK Data Service (registration required).

UK Data Service: <https://ukdataservice.ac.uk/>

#### 4.3 Understanding Society (UKHLS)

**What it contains:**

Longitudinal household survey continuing the British Household Panel Survey.  
Includes wealth modules with inheritance data.

**Access:**

URL: <https://www.understandingsociety.ac.uk/>

Free registration required. Special license for sensitive variables.

**Key Variables:**

- w\_inherit: Whether received inheritance
- w\_inheramt: Amount inherited
- w\_inherwho: Relationship to deceased

#### 5. INTERNATIONAL COMPARISON DATA

##### 5.1 OECD Data

**Inheritance/Estate/Gift Tax Revenue:**

URL: <https://stats.oecd.org/> → Tax → Revenue Statistics

Variable: Tax on property → Recurrent taxes on net wealth → Estate, inheritance and gift taxes (4300)

**Wealth Distribution:**

URL: <https://stats.oecd.org/> → Wealth Distribution Database

Contains: Gini coefficients, wealth shares by percentile, for 25+ countries.

##### 5.2 Luxembourg Wealth Study (LWS)

**What it contains:**

Harmonized microdata on household wealth from 15+ countries. Enables cross-country comparison with consistent definitions.

**Access:**

URL: <https://www.lisdatacenter.org/our-data/lws-database/>

Free for academic researchers. Remote execution system (no direct data download).

### 5.3 World Inequality Database

#### **What it contains:**

Long-run data on income and wealth inequality for most countries. Includes top wealth shares from estate tax data where available.

#### **Access:**

URL: <https://wid.world/>

Free download of all data and methodology documentation.

## 6. STEP-BY-STEP ANALYSIS GUIDE

### 6.1 Structural Estimation: Bequest Elasticity

#### **Step 1: Compile estate tax policy variation**

- Create state-year panel of estate/inheritance tax parameters (1980-2025)
- Note federal exemption changes: 1997 (\$600K), 2001 (\$675K), 2002-2009 (phased to \$3.5M), 2010 (repealed), 2011-2012 (\$5M), 2013-2017 (\$5.25M indexed), 2018+ (\$11M+)

#### **Step 2: Merge with estate tax return data**

- Use IRS SOI public-use files for state-year aggregates
- Restricted access provides individual-level data for finer analysis

#### **Step 3: Estimate elasticity**

Specification (state-year level):

$$\ln(\text{estates}_{st}) = \beta \cdot \ln(1 - \tau_{st}) + X_{st}'\gamma + \mu_s + \delta_t + \varepsilon_{st}$$

where  $\tau_{st}$  is the combined federal-state marginal tax rate.

The coefficient  $\beta$  estimates the elasticity of reported estates with respect to the net-of-tax rate.

#### **Step 4: Robustness checks**

- Bunching analysis around exemption thresholds
- Event studies around major tax changes (EGTRRA 2001, TCJA 2017)
- Placebo tests using non-taxable estates

### 6.2 Cross-Country Analysis: UK vs. US

#### **Step 1: Construct comparable samples**

- US: SCF households with wealth > \$1M

- UK: WAS households with wealth > £800K (PPP-adjusted)
- Match on demographics, income, wealth level

### **Step 2: Estimate outcome differences**

$$y_{it} = \alpha + \beta \cdot \text{UK}_i + X_{it}'\gamma + \delta_t + \varepsilon_{it}$$

Outcomes: number of beneficiaries, concentration of bequests, charitable giving, effective tax rate.

### **Step 3: Within-country time series**

- Examine changes around UK IHT reforms (2006 transferable nil-rate band, 2017 residence nil-rate band)
- Compare to US changes in same period

#### 6.3 State-Level Diff-in-Diff

### **Step 1: Define treatment and control**

Treatment: States that repealed inheritance tax after 2001

Control: States that retained inheritance tax (IA, KY, MD, NE, NJ, PA)

### **Step 2: Construct outcome variables**

- Bequest concentration (share to top beneficiary)
- Number of beneficiaries per estate
- Charitable bequests as share of estate
- Wealth migration (in/out of state)

### **Step 3: Estimate effect**

$$y_{ist} = \alpha + \beta \cdot (\text{Post}_t \times \text{Repeal}_s) + \mu_s + \delta_t + X_{ist}'\gamma + \varepsilon_{ist}$$

### **Step 4: Event study specification**

$$y_{ist} = \alpha + \sum_k \beta_k \cdot (\text{Year}_t = k) \times \text{Repeal}_s + \mu_s + \delta_t + \varepsilon_{ist}$$

Plot  $\beta_k$  coefficients to visualize pre-trends and treatment effects.

## 7. SAMPLE CODE TEMPLATES

### 7.1 Stata: Bequest Elasticity Estimation

\* Load IRS SOI data

```
use soi_estate_tax_panel.dta, clear
```

\* Generate variables

```
gen ln_estates = ln(num_estates)
```

```
gen net_of_tax = 1 - marginal_rate
```

```

gen ln_net_of_tax = ln(net_of_tax)
* Basic elasticity regression
reghdfe ln_estates ln_net_of_tax gdp_per_capita pop_over65, ///
absorb(state year) cluster(state)
* Store coefficient
local elasticity = _b[ln_net_of_tax]
di "Bequest elasticity: " `elasticity'
7.2 Stata: State-Level Diff-in-Diff
* Load merged state-level data
use state_inheritance_panel.dta, clear
* Define treatment
gen repeal_state = inlist(state, "IN", "OH", "NC", ...) // states that repealed
gen post = (year >= 2002)
gen treat = repeal_state * post
* Diff-in-diff
reghdfe bequest_concentration treat, absorb(state year) cluster(state)
* Event study
gen rel_year = year - 2001
forvalues k = -5/10 {
    gen rel_`k' = (rel_year == `k') * repeal_state
}
reghdfe bequest_concentration rel_* (omit: rel_-1), absorb(state year) cluster(state)
coefplot, vertical
7.3 R: Cross-Country Analysis
# Load packages
library(tidyverse)
library(fixest)
library(haven)
# Load US SCF data
scf <- read_dta("scf2022.dta") %>%

```

```

filter(networth > 1000000) %>%
mutate(country = "US")
# Load UK WAS data
was <- read_dta("was_wave7.dta") %>%
filter(totwlth > 800000) %>%
mutate(country = "UK")
# Combine and estimate
combined <- bind_rows(scf, was)
# Regression
model <- feols(num_beneficiaries ~ UK + age + income + wealth | year,
data = combined, cluster = ~country_year)
summary(model)

```

## 8. PROJECT TIMELINE AND CHECKLIST

### 8.1 Phase 1: Data Assembly (Weeks 1-8)

- Download IRS SOI public-use estate tax files
- Download SCF data (2019, 2022 waves)
- Register for PSID access and create custom extract
- Download HMRC inheritance tax statistics
- Register for UK Data Service and download WAS
- Compile state inheritance tax parameters (1990-2025)
- Download Census state government finance data
- Download OECD tax revenue and wealth distribution data
- Submit application for IRS restricted data access (if pursuing)

### 8.2 Phase 2: Data Cleaning (Weeks 9-16)

- Harmonize US and UK wealth definitions
- Create state-year panel of tax parameters
- Link estate tax data to state characteristics
- Construct inheritance receipt variables from SCF/PSID
- Create consistent beneficiary classification across datasets
- Document all variable definitions and coding decisions

### 8.3 Phase 3: Estimation (Weeks 17-28)

- Estimate bequest elasticity (Section 6.1)
  - Estimate planning effectiveness parameter
  - Estimate dispersion preference parameter
  - Run cross-country UK vs. US analysis (Section 6.2)
  - Run state-level diff-in-diff (Section 6.3)
  - Conduct robustness checks and sensitivity analysis
  - Bootstrap standard errors
- 8.4 Phase 4: Write-Up (Weeks 29-36)
- Update paper with empirical results
  - Create publication-quality tables and figures
  - Write detailed data appendix
  - Prepare replication package
  - Submit to journal

## 9. KEY CONTACTS AND RESOURCES

### 9.1 Data Access Contacts

#### **IRS Statistics of Income:**

Email: sis@irs.gov

Phone: (202) 317-3100

For research proposals: SOI.Research@irs.gov

#### **Federal Statistical Research Data Centers:**

URL: <https://www.census.gov/about/adrm/fsrdc.html>

Provides access to restricted IRS, Census, and other federal microdata.

#### **UK Data Service:**

Email: help@ukdataservice.ac.uk

Phone: +44 (0)1206 872143

### 9.2 Expert Contacts

Researchers with expertise in estate/inheritance taxation who have used relevant data:

- Wojciech Kopczuk (Columbia) — Estate tax behavioral responses
- David Joulfaian (Treasury, retired) — Estate tax data
- Lily Batchelder (NYU Law) — Inheritance tax policy

- Emmanuel Saez (Berkeley) — Wealth inequality
- Gabriel Zucman (Berkeley) — Wealth data and taxation

### 9.3 Software Requirements

- Stata 17+ (for reghdfe, coefplot packages)
- R 4.0+ with tidyverse, fixest, haven packages
- Python 3.8+ with pandas, statsmodels (optional)
- LaTeX for paper compilation

### 9.4 Estimated Costs

<b>Item</b>	<b>Cost</b>
IRS SOI public data	Free
SCF, PSID data	Free
UK Data Service access	Free (academic)
FSRDC access (if needed)	\$20,000-50,000/year
CCH State Tax Handbook (historical)	~\$200/year
Stata license	~\$600 (academic)
Research assistant (optional)	\$15-25/hour

Note: Most analyses can be completed with freely available public data. FSRDC access is only necessary for individual-level estate tax return analysis and is optional for the core findings.