

Flexible Subsetting and Visualization for Policy Data

A Stata Workflow for a Faster Path to Research-Ready Data

Andrea Chu

PhD Candidate in Economics
American University

2026 Stata Conference
October 2, 2026

The Big Picture

Data Challenge

Data constructed for yesterday's questions may not be structured for today's research questions.

The Big Picture

Data Challenge

Data constructed for yesterday's questions may not be structured for today's research questions.

Today's Proposal: A Faster Path to Research-Ready Data

This presentation introduces a **flexible Stata workflow** that preserves the details first and simplifies them later.

Table of Contents

1 Motivation: Data Challenges

- Visualization 1: From Categories to Overlapping Sets

2 Proposal: A Flexible Stata Workflow

- Visualization 2: From Linked Attributes to Case Identification

3 Conclusion

Table of Contents

1 Motivation: Data Challenges

- Visualization 1: From Categories to Overlapping Sets

2 Proposal: A Flexible Stata Workflow

- Visualization 2: From Linked Attributes to Case Identification

3 Conclusion

Motivation: Data Challenges

How can researchers preserve policy complexity while making sense of the data?

- **Multidimensional policy actions** can span channels, sectors, entities, and implementation stages.

Motivation: Data Challenges

How can researchers preserve policy complexity while making sense of the data?

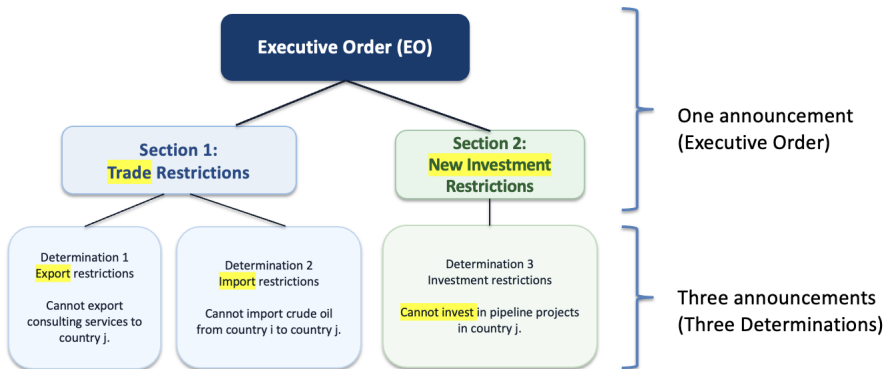
- **Multidimensional policy actions** can span channels, sectors, entities, and implementation stages.
- **Fragmented records** can obscure how individual measures connect to a broader policy theme.

Motivation: Data Challenges

How can researchers preserve policy complexity while making sense of the data?

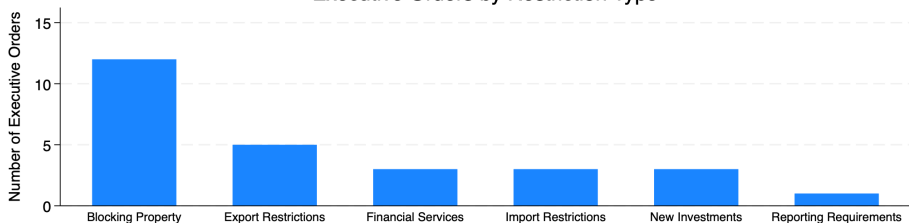
- **Multidimensional policy actions** can span channels, sectors, entities, and implementation stages.
- **Fragmented records** can obscure how individual measures connect to a broader policy theme.
- **Evolving analytical lens:** As the literature, policy context, and researchers' understanding evolve, existing classifications may need to be refined or redefined.

Structural View of Economic Policy

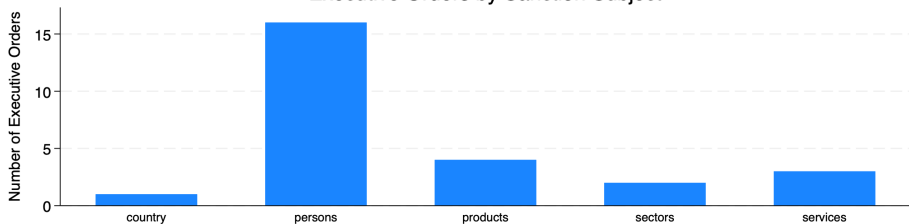


Separate Views of Policy Attributes

Executive Orders by Restriction Type



Executive Orders by Sanction Subject



Visualizing Overlapping Policy Attributes

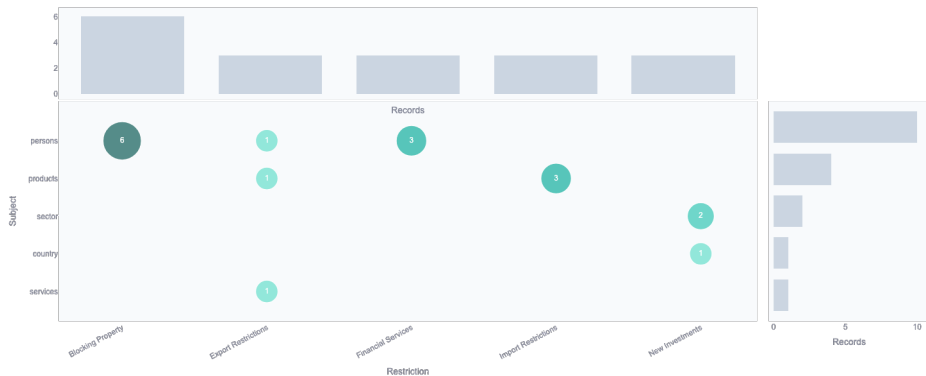


Table of Contents

1 Motivation: Data Challenges

- Visualization 1: From Categories to Overlapping Sets

2 Proposal: A Flexible Stata Workflow

- Visualization 2: From Linked Attributes to Case Identification

3 Conclusion

Proposal: A Faster Path to Research-Ready Data

A Simple and Efficient Workflow

- 1 **Information Capture:** Capture policy attributes in a flexible format.
- 2 **Preview:** Preview available attributes as research questions evolve.
- 3 **Data Extraction:** Select relevant attributes and convert them into analysis-ready columns.

Step 1: Information Capture

Presidential Documents: Executive Order 14024 of April 15, 2021

Section 1. All property and interests in property that are in the United States, that hereafter come within the United States, or that are or hereafter come within the possession or control of any United States person of the following persons are blocked and may not be transferred, paid, exported, withdrawn, or otherwise dealt in:

(a) any person determined by the Secretary of the Treasury, in consultation with the Secretary of State, and, with respect to subsection (a)(ii) of this section, in consultation with the Attorney General, or by the Secretary of State, in consultation with the Secretary of the Treasury, and, with respect to subsection (a)(ii) of this section, in consultation with the Attorney General:

1(a)(i) (i) to operate or have operated in the technology sector or the defense and related materiel sector of the Russian Federation economy, or any other sector of the Russian Federation economy as may be determined by the Secretary of the Treasury, in consultation with the Secretary of State;

1(a)(ii) (ii) to be responsible for or complicit in, or to have directly or indirectly engaged or attempted to engage in, any of the following for or on behalf of, or for the benefit of, directly or indirectly, the Government of the Russian Federation:

(A) malicious cyber-enabled activities;

(B) interference in a United States or other foreign government election;



Data Editor (Browse)

EO_number[1]

14024

DocID	EO_number	Note1	Note2
1	1	14024	

Section 1(a)(i)

Section 1(a)(ii)

EO_section="1(a)(i);restriction="Blocking Property";sanction_subject="pers...

EO_section="1(a)(ii);restriction="Blocking Property";sanction_subject="per...

Information is stored as key–value pairs separated by semicolons.

Step 2: Preview

i. Raw data

Data Editor (Browse)					Edit mode	
DocID[1]		1				
	DocID	EO_number	Note1	Note2		
1	1	14024	ref="1(a)(i)";restriction="Blocking Property";sanction_subject="persons";channel="finance";...	ref="1(a)(ii)";restriction="Blocking Property";sanction_subject="persons";channel="finance";sanction_...		
2	2	14039	restriction="Blocking Property";ref="1(a)";sanction_subject="persons";sanction_scope="Ce...	restriction="Blocking Property";ref="3(a)";sanction_subject="persons";sanction_scope="Certain Russ		
3	3	14066	ref="1a)(i)";restriction="Certain Imports";sanction_subject="products";sanction_subject_desc...	ref="1a)(ii)";restriction="New Investments";sanction_subject="sector";sanction_subject_desc="energy		
4	4	14068	restriction="Certain Imports";ref="1(a)(i)";sanction_subject="products";sanction_subject_desc...	restriction="Certain Exports";ref="1(a)(ii)";sanction_subject="products";sanction_subject_desc="luxur.		
5	5	14071	restriction="New Investments";sanction_subject="country";sanction_subject_desc="Russia"...	restriction="Certain Exports";ref="1(a)(ii)";sanction_subject="services";sanction_subject_desc="any c.		
6	6	14114	restriction="Blocking Property";parent_EO=14024;note="adding a new section 11 to EO 14...	parent_EO=14068;restriction="import restrictions";ref="3";note="by striking sub1(a)(i) and inserting";s.		

ii. Command

* tagpreview reveals the available attributes ready for extraction

```
tagpreview Note1 Note2 Note3 Note4 Note5 // Columns containing the information to be extracted
```



After transposing entries

Total entries: 19

Attribute	Count	Percentage
EO_section	19	100%
restriction	19	100%
sanction_subject	18	95%
sanction_subject_desc	15	79%

Step 3: Data Extraction

* taggenerate turns selected attributes from the preview into an analysis-ready dataset.

```
taggenerate EO_section restriction sanction_subject  
    sanction_subject_desc sanction_actor  
    sanction_scope sanction_scope_desc           // (a)select attributes identified in the preview  
    , columns(Note1 Note2 Note3 Note4 Note5)      // (b)specify source columns containing the attributes  
    keep(DocID EO_number Signed_date Effective_date) // (c)retain these columns unchanged  
    saving("data_EO.dta")                        // (d)specify the name of the newly created dataset
```

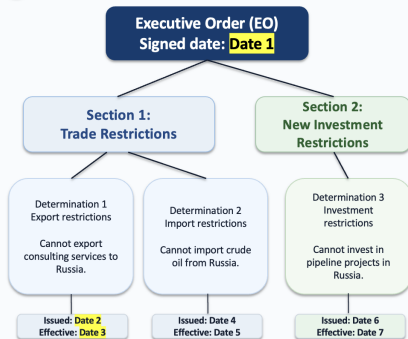


Data Editor (Browse) — data_EO.dta

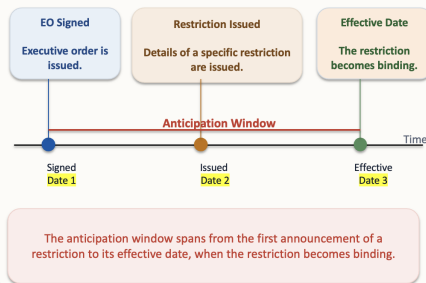
	DocID	EO_number	EO_section	restriction	sanction_subject	sanction_subject_desc	sanction_actor	sanction_scope
1	1	14024	1(a)(i)	Blocking Property	persons			sector
2	1	14024	1(a)(ii)	Blocking Property	persons			activities
3	2	14039	1(a)	Blocking Property	persons	any foreign person identified pursu...		Certain Russian EnergyExport Pipelines
4	2	14039	3(a)	Blocking Property	persons			Certain Russian EnergyExport Pipelines
5	2	14039	3(b)	Blocking Property	persons	the receipt of any contribution or pr...		Certain Russian EnergyExport Pipelines
6	3	14066	1a)(i)	Certain Imports	products	crude oil, petroleum, petroleum fue...		
7	3	14066	1a)(ii)	New Investments	sector	energy sector	United States person	
8	3	14066	1a)(iii)	Financial Services	person	energy sector	United States person	
9	4	14068	1(a)(i)	Certain Imports	products	fish, seafood, and preparations the...		
10	4	14068	1(a)(ii)	Certain Exports	products	luxury goods		
11	4	14068	1(a)(iii)	New Investments	sector	any sector to be determined		
12	4	14068	1(a)(iv)	Certain Exports	person	U.S. dollar-denominated banknotes	Rus gov or any person in Rus	
13	4	14068	1(a)(v)	financial services	person	covered in section(1)	United States person	
14	5	14071	1(a)(i)	New Investments	country	Russia		
15	5	14071	1(a)(ii)	Certain Exports	services	any category of services to be det...		
16	5	14071	1(a)(iii)	financial services	person	Russia	United States person	
17	6	14114	1	Blocking Property	persons	operating or having operated in th...	foreign financial institution	
18	6	14114	3	import restrictions	products	fish, seafood, and preparations ther...		
19	6	14114	3	reporting requirement for enforcement				

Two Complementary Views of Economic Policy

1 Structural View: Policy Hierarchy



2 Temporal View: From Announcement to Implementation



Import & Export Restrictions: Shorter Anticipation Windows

Dimension ⓘ

Sector

Restrictions

☐ Select all ⓘ

Highlight values ⓘ

export restrictions

blocking property

import restriction

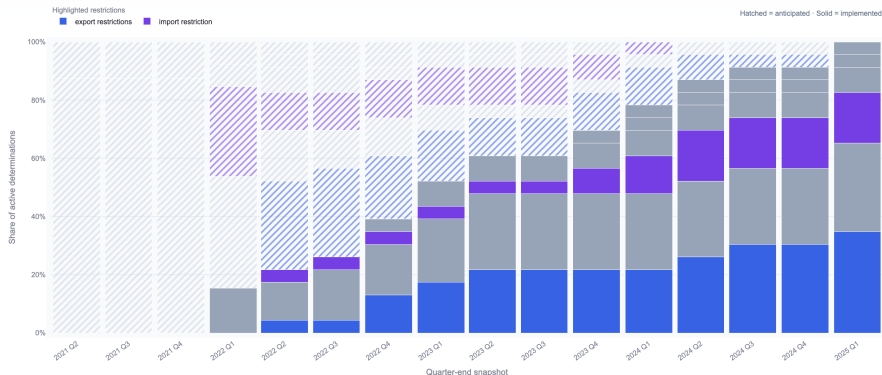
price cap policy

facilitate transactions

import restriction; diamond jewelry and unsorted diamonds

2 of 6 values highlighted

Status is measured at quarter end. Hatched segments are anticipated; solid segments are implemented.



Price Cap Policy: Longer Anticipation Window

Dimension ⓘ

Sector

Restrictions

☐ Select all ⓘ

Highlight values ⓘ

export restrictions

blocking property

import restriction

price cap policy

facilitate transactions

import restriction; diamond jewelry and unsorted diamonds

1 of 6 values highlighted

Status is measured at quarter end. Hatched segments are anticipated; solid segments are implemented.

Highlighted restrictions

■ price cap policy

Hatched = anticipated · Solid = implemented

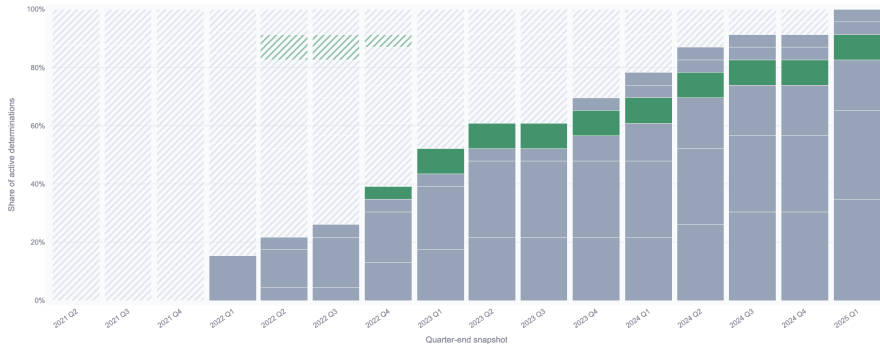


Table of Contents

1 Motivation: Data Challenges

- Visualization 1: From Categories to Overlapping Sets

2 Proposal: A Flexible Stata Workflow

- Visualization 2: From Linked Attributes to Case Identification

3 Conclusion

Takeaway: A Faster Path to Research-Ready Data

Data Challenge

Data constructed for yesterday's questions may not be structured for today's research questions.

Takeaway: A Faster Path to Research-Ready Data

Data Challenge

Data constructed for yesterday's questions may not be structured for today's research questions.

A Flexible Workflow

Build and refine the dataset as research questions and understanding evolve.

- 1 **Information Capture**
- 2 **Preview**
- 3 **Data Extraction**

Takeaway: A Faster Path to Research-Ready Data

Data Challenge

Data constructed for yesterday's questions may not be structured for today's research questions.

A Flexible Workflow

Build and refine the dataset as research questions and understanding evolve.

- 1 **Information Capture**
- 2 **Preview**
- 3 **Data Extraction**

Code and Examples



github.com/andreachu-research/entrytags
Andrea Chu: tc1355a@american.edu