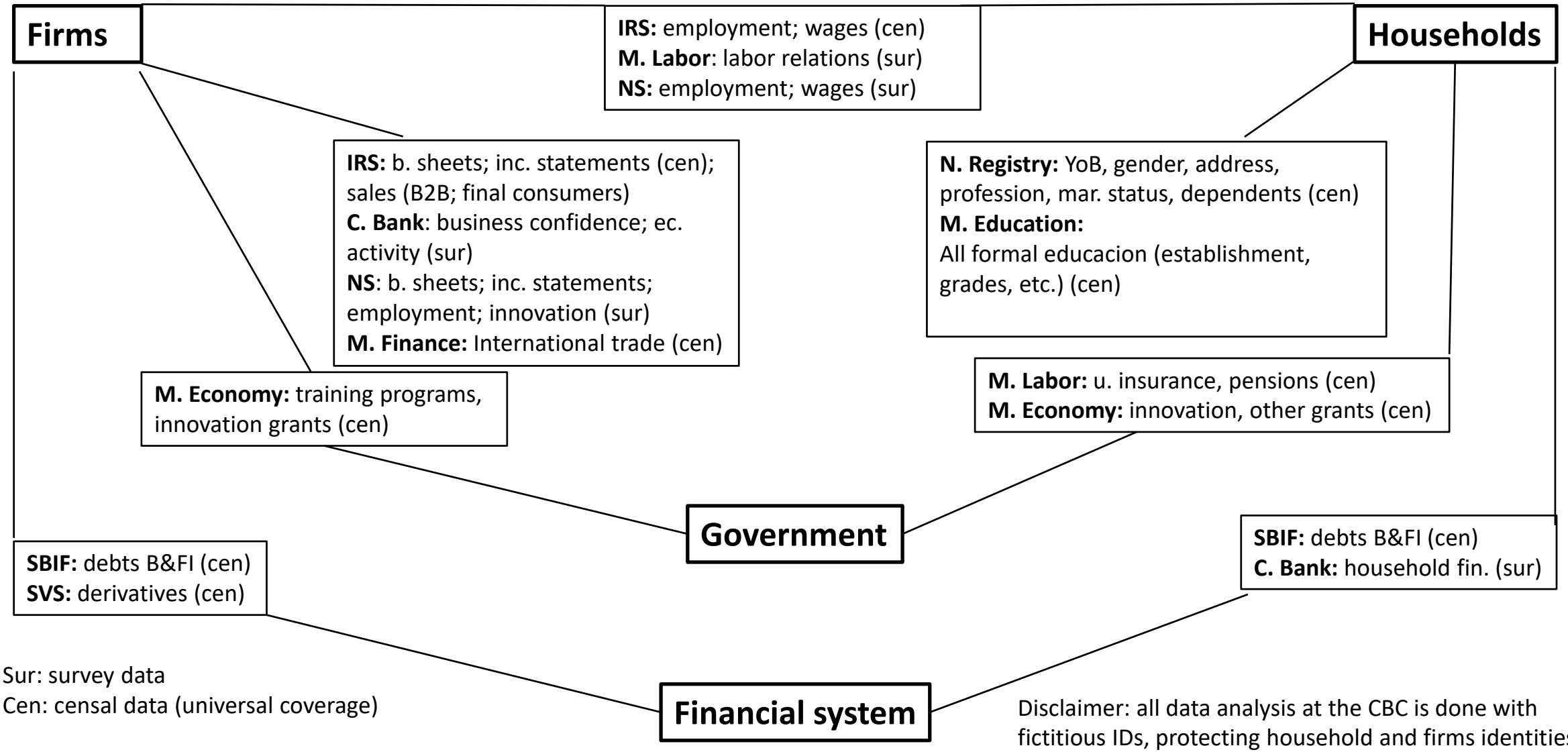
Micro data in Central Banks: New tools for Macro Analysis

Elias Albagli
Director of Monetary Policy Division & Chief Economist

Central Bank of Chile
November 2020

Chilean administrative data is very good by international standards, but it remains *lonely*.
Our effort at CBC: merge separate datasets for macro and applied policy analysis exploiting existence of unique identifiers at firm and household level



Sur: survey data
Cen: censal data (universal coverage)

Disclaimer: all data analysis at the CBC is done with fictitious IDs, protecting household and firms identities.

Agenda

Benefits of micro data

1. Speed
2. Depth
3. Policy

Obstacles

4. Cooperation and perception

Agenda

Benefits of micro data

1. **Speed**
2. Depth
3. Policy

Obstacles

4. Cooperation and perception

Usual macro stats (activity, demand, labor mkt, inflation): 1-3 months lag. Compare this with admin. data lags:

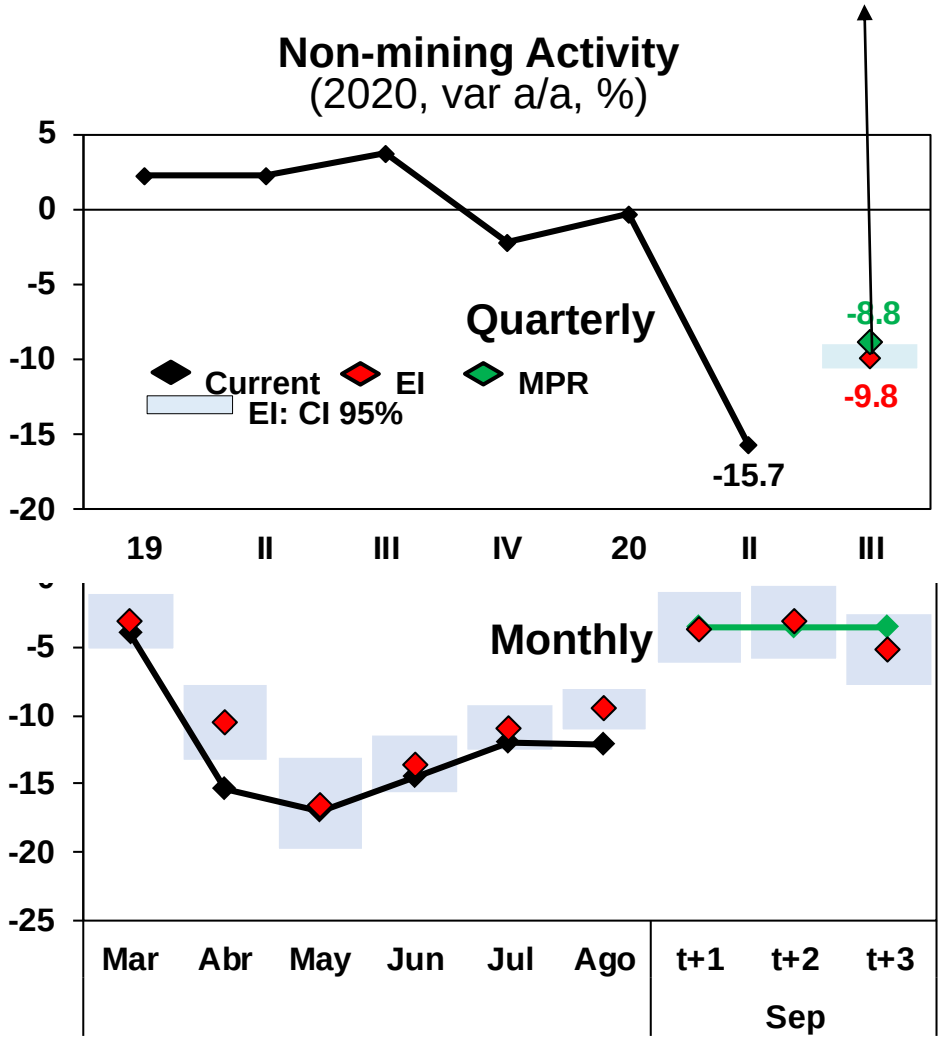
- **Activity:**
  - IRS E-invoice (factura electrónica)
    - All B2B transactions between Chilean firms (in the formal sector) – **daily, one week lag**
    - Includes: goods/services description; unit Price; quantity; discounts; delivery address
- **Aggregate demand:**
  - Consumption:
    - E-receipts (boleta electrónica): B2C transactions of large Chilean retailers; starting Jan. 2021: universe of retailers –**daily, one week lag**
    - Electronic payments: universe of B2C and B2B transactions paid through debit and credit cards –**daily, one month lag**
  - Investment:
    - VAT declarations: **monthly frequency, one month lag**; B2B transactions: **daily, one week lag**
  - International trade;
    - Customs declarations: **daily, one-week lag**
- **Labor market**
  - Social sec. & unemployment insurance contributions; contract termination declarations – **monthly, one month lag**
  - Includes: firm-worker pair information (tenure; wages; reason for contract termination)
- **Inflation**
  - (Firm-specific) producer price indexes: **daily, one week lag**
  - (Firm-specific) consumer price indexes (starting Jan 2021): **daily, one week lag**
  - (Firm-specific) wage costs: **monthly, one month lag**

Our latest application:
forecasting recovery from the bottomless pit

Electronic Invoice
(var y/y, %)

Sector	Septiembre		20.T3			
	FE t+3	IPoM Sep	FE t+3	inf	sup	IPoM Sep
Industria	6.3%	4.0%	-3.1%	-4.0%	-2.2%	-2.9%
Construcción	-20.5%	-13.0%	-28.3%	-29.6%	-26.8%	-26.4%
Comercio y RR&HH	-3.1%	-5.3%	-10.1%	-12.3%	-7.7%	-13.6%
Transporte	-19.9%	-4.8%	-21.2%	-22.4%	-20.0%	-13.1%
Serv. Empresariales	-8.1%	-3.0%	-10.4%	-12.0%	-8.7%	-6.7%
Serv. Personales	-24.0%	-13.0%	-19.1%	-24.9%	-11.7%	-14.6%
No Minero	**-5.2%**	**-3.6%**	**-9.8%**	**-10.7%**	**-8.9%**	**-8.8%**

Q3 effective: -10%



Note: **'t+1', 't+2', and 't+3' are succesive forecast after weekly arrival of new information.**

Agenda

Benefits of micro data

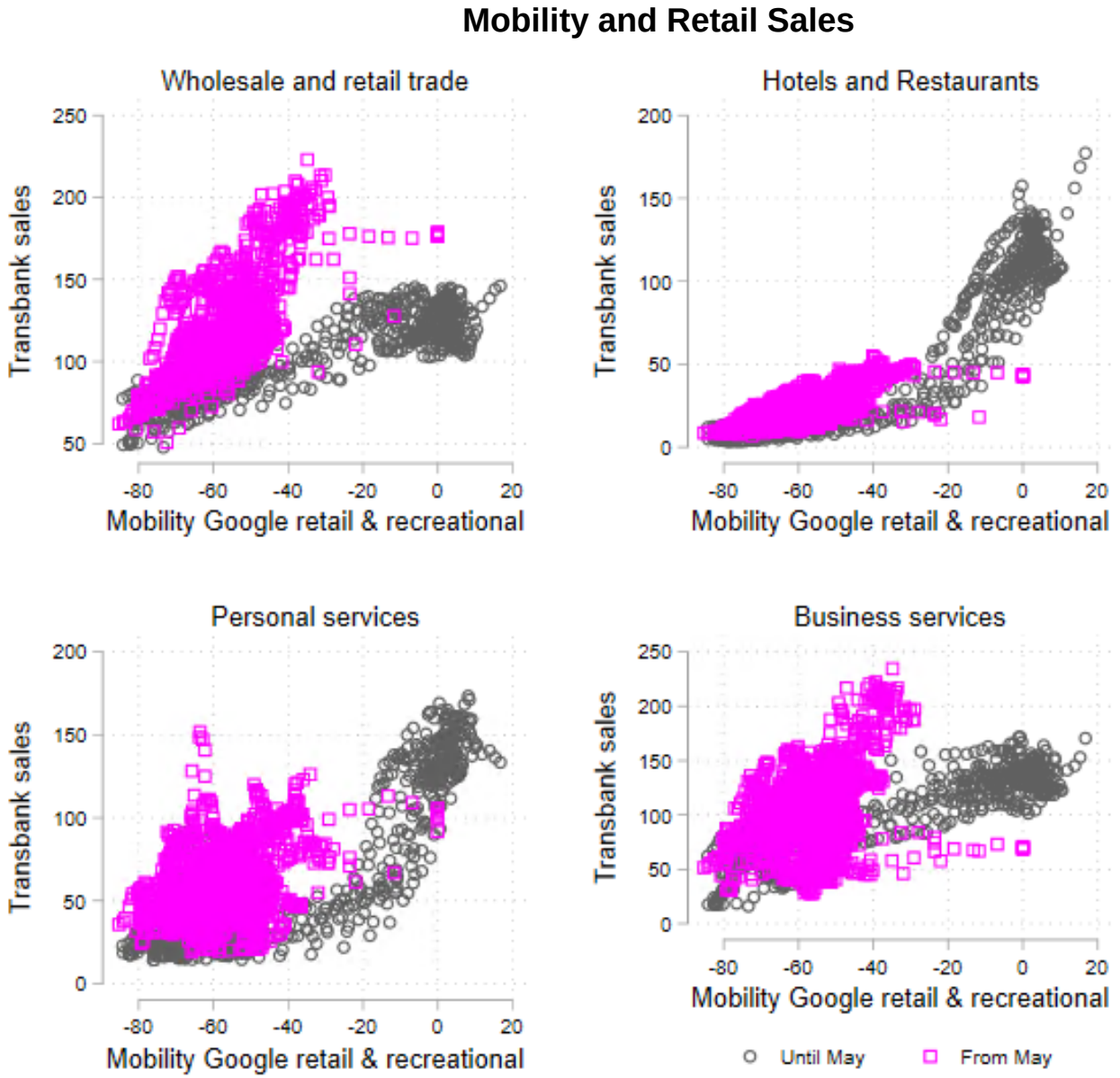
1. Speed
2. **Depth**
3. Policy

Obstacles

4. Cooperation and perception

While speed is useful, depth (disaggregation) is even better

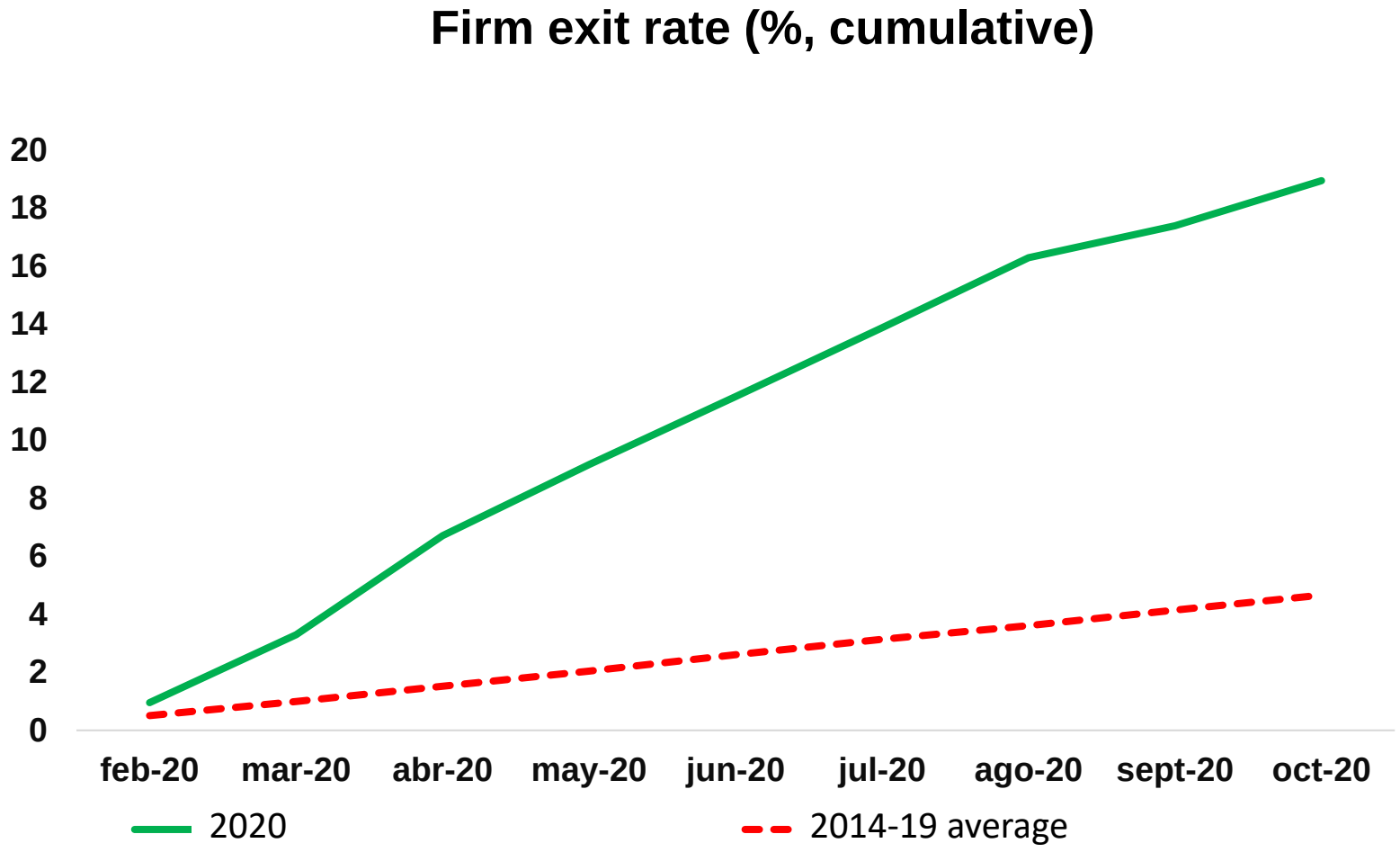
1. **How does mobility affect sales, for different sectors?**
   - Daily sales (by region/sector) and google mobility index (2019=100).
   - Between March-May: strong link between sales and mobility in most sectors.
   - After May: retail and business services increase resilience, but not Hotels&Restaurants, Personal Services.
   - ***Insight:*** conditional on restricted mobility, scenario might have better performance in some sectors, but others have little/no margin of adjustment.



While speed is useful, depth (disaggregation) is even better

2. How persistent will the scarring effects of the crisis be?

- Aggregate macro stats don´t allow an evaluation of firm death rates –even ex post, bankruptcy data is heavily skewed towards larger firms, with smaller establishments dying “off the radar”.
- However, this seems crucial to evaluate scarring effects, and calibrate medium term projections beyond lockdowns
- ***Insight***: potential output has been depressed throughout the pandemic, and its recovery may be slow due to the permanent exit of a significant number of firms.



Agenda

Benefits of micro data

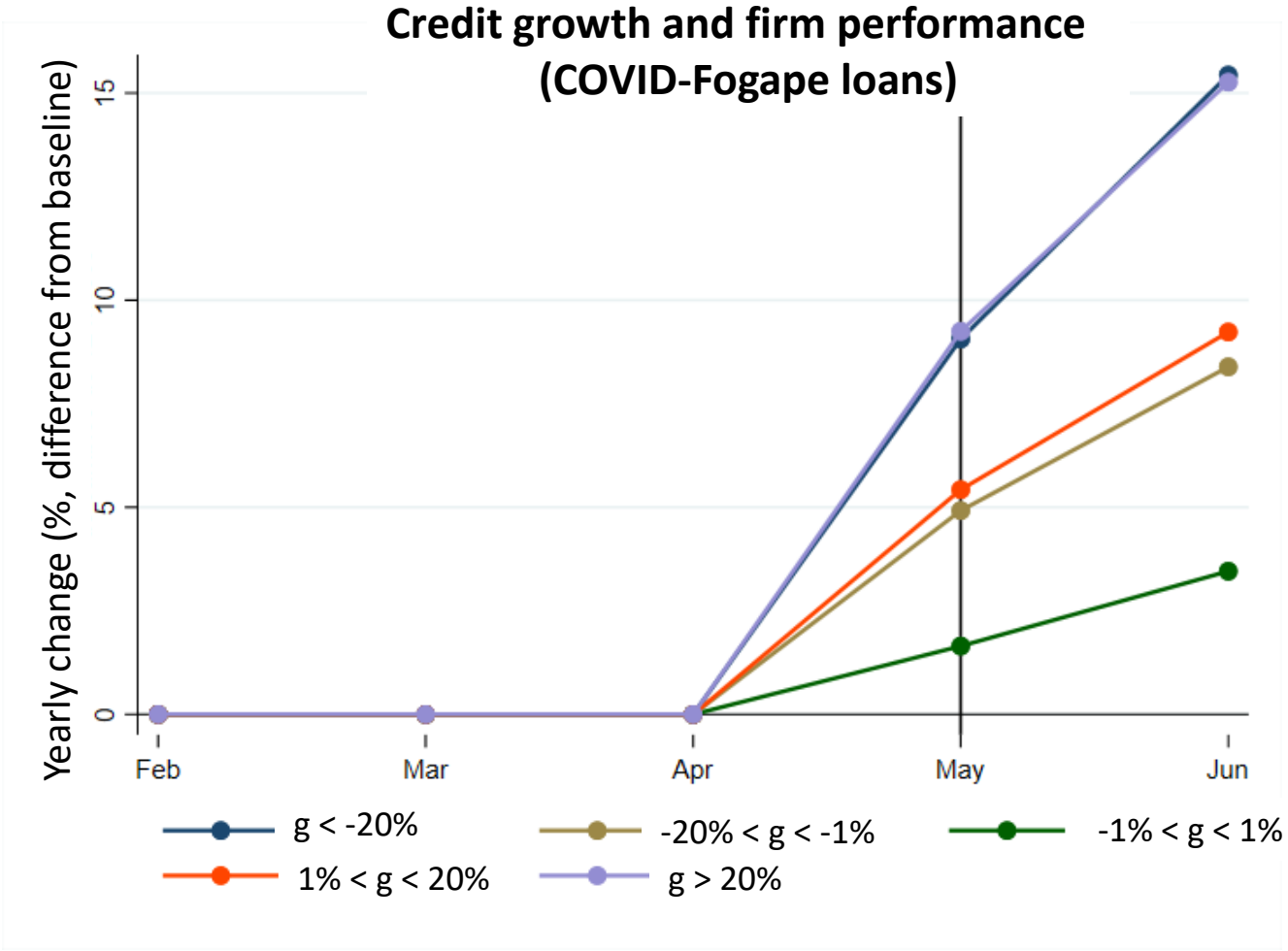
1. Speed
2. Depth
3. **Policy**

Obstacles

4. Cooperation and perception

When it comes to policy evaluation, merging is the key

1. **Have new credit facilities at the CBC (similar to TLTROs) mobilized credit where it is most needed?**
   - Aggregate credit data tells us that, contrary to other recessions, credit behaved countercyclically.
   - But does this reflect large firms hoarding cheap credit? Or is credit flowing to SMEs?
   - ***Insight:*** the program exceeded expectations → growth has been the largest where its most needed (and also where it finances growth).



When it comes to policy evaluation, merging is the key

2. **Has credit access made a difference?**
   - Did access to credit affect real firm decisions, such as survival; investment; employment?
   - ***Insight:*** access to credit negatively related to exit probability (report 0 sales) and positively related to investment performance.
   - (Effects on employment? We just received employment data!)

$$Y_{i,T} = \beta 0 + \beta 1 * C_{i,T-1} + \beta 2 * V_{i,T-1} + \beta 3 * V_{i,T-1} * C_{i,T-1} + \beta 4 * F_i + e_i$$

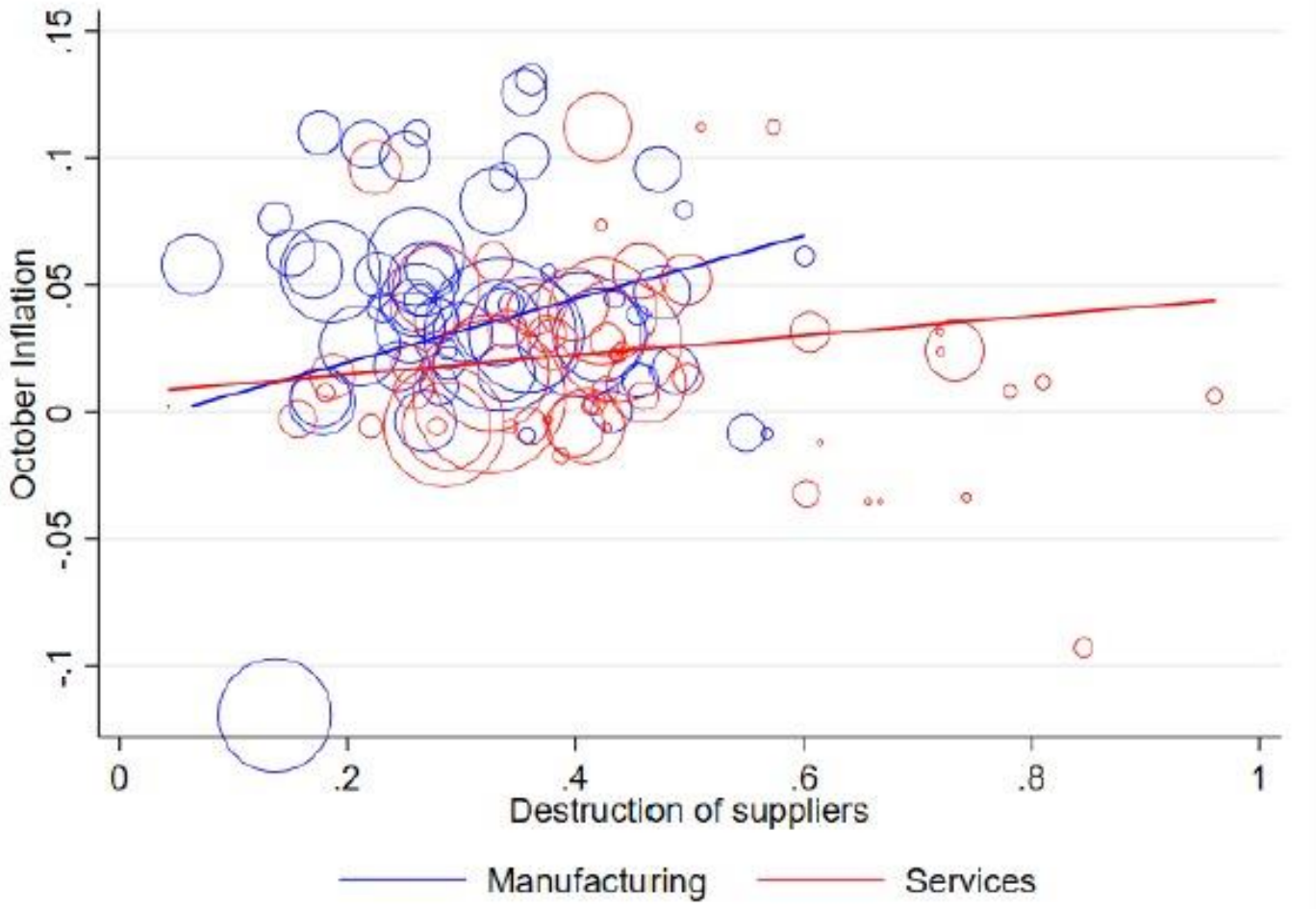
	(1)	(2)
	No reporte	No reporte
Crecimiento ventas	-0.0363***	-0.0386***
	(0.00203)	(0.00227)
Crecimiento deuda	-0.00444***	
	(0.00109)	
Crecimiento ventas * Crecimiento deuda	0.00115	
	(0.00151)	
Deuda nueva=1		-0.0198***
		(0.00267)
Deuda nueva=1 * Crecimiento ventas		0.00855**
		(0.00382)
Constante	0.0787***	0.0890***
	(0.00137)	(0.00158)
Tamaño	Si	Si
Sector	Si	Si
Edad	Si	Si
Observaciones	41874	46483
R^2	0.048	0.048
R^2 Ajustado	0.048	0.048

Nota: Errores estándares en paréntesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

When it comes to policy evaluation, merging is the key

3. Is our recent (positive) inflation surprise transitory, or persistent?

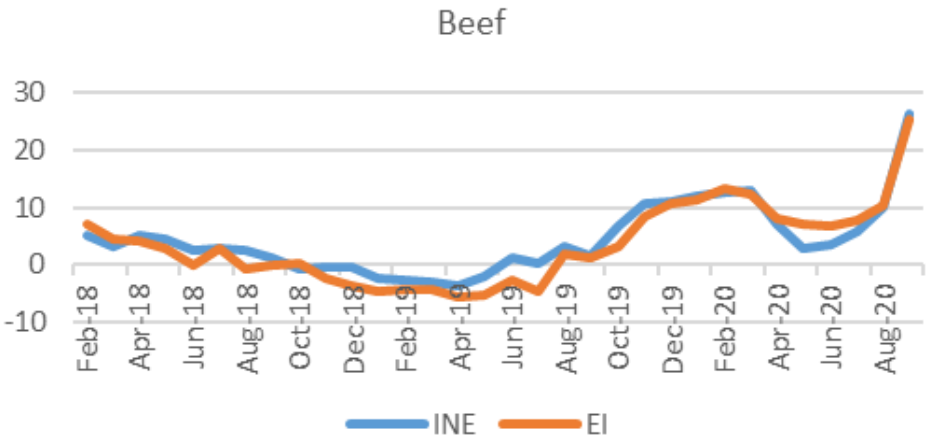
- CPI in August-October surprised significantly to the upside: 76 bp (target: 3%).
- This likely reflects consumption increased due to the withdrawal of 10% of private pensions savings, which should be transitory (and not affect inflation at the MP horizon of 2 years).
- But… might there be something more to it? Perhaps more persistent effects from supply-side disruptions? Data from e-invoicing includes identities of seller and buyer – allows tracing changes in supplier networks.
- ***Insight***: there seems to be a positive relation between supply-chain disruptions and prices --warrants further analysis (preliminary!)

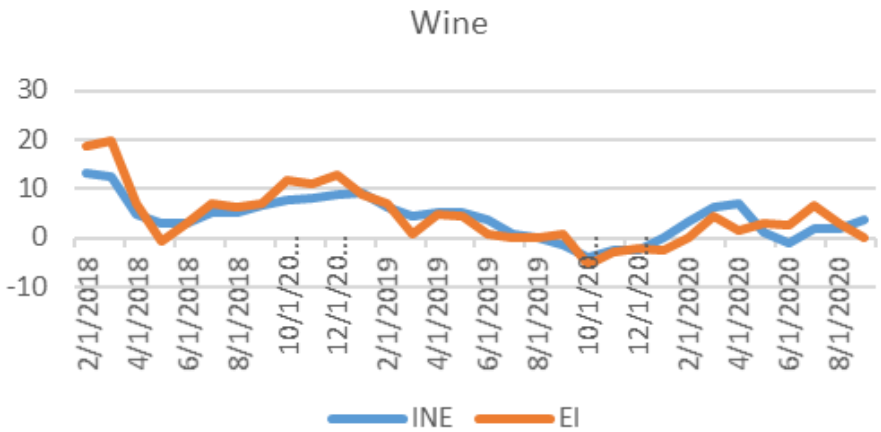


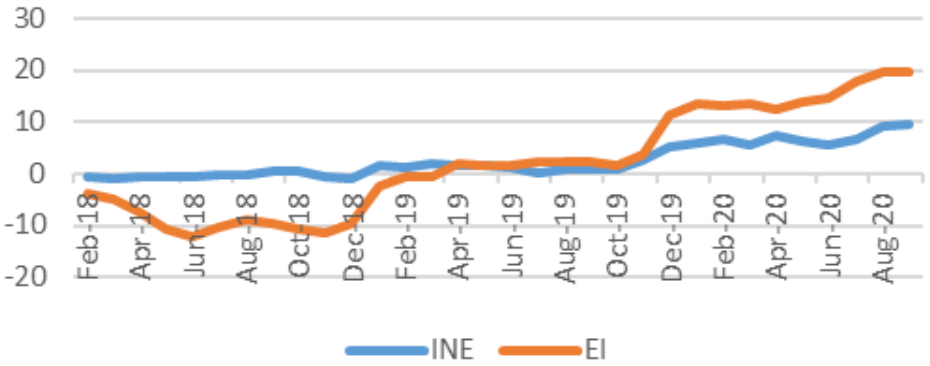
Next steps: understanding inflation pressures at the micro level

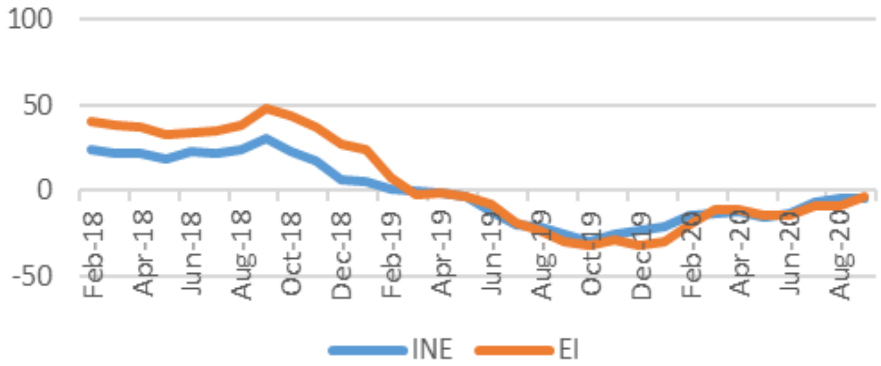
- **E-invoicing (factura electronica) records prices and quantities separately**
  - We can construct price indexes for the goods/services sold by each firm (for B2B transactions).
  - This implies we also observe the cost of purchased materials, electricity bills, etc.
  - Merging with int. trade data, we also observe costs of imported inputs.
  - We also observe wages at the firm level.
  - Of course, this all requires time and resource-intensive machine learning algorithms to codify information from non-standardized reporting (goods and services description).
- **E-receipts (boleta electronica)**
  - Starting January 2021, we will also receive individual B2C sales, for all firm in the country (exponentially larger datasets). This also includes prices and quantities separately.
  - This would lead to a direct measure of CPI inflation, weighted by the effective expenditure basket.
- Together, these price-related datasets will allow us to estimate firm-level Phillips curve-type relationships, improving our forecasting capacity and the calibration of MP decisions.

A small sample of recent progress…









Agenda

Benefits of micro data

1. Speed
2. Depth
3. Policy

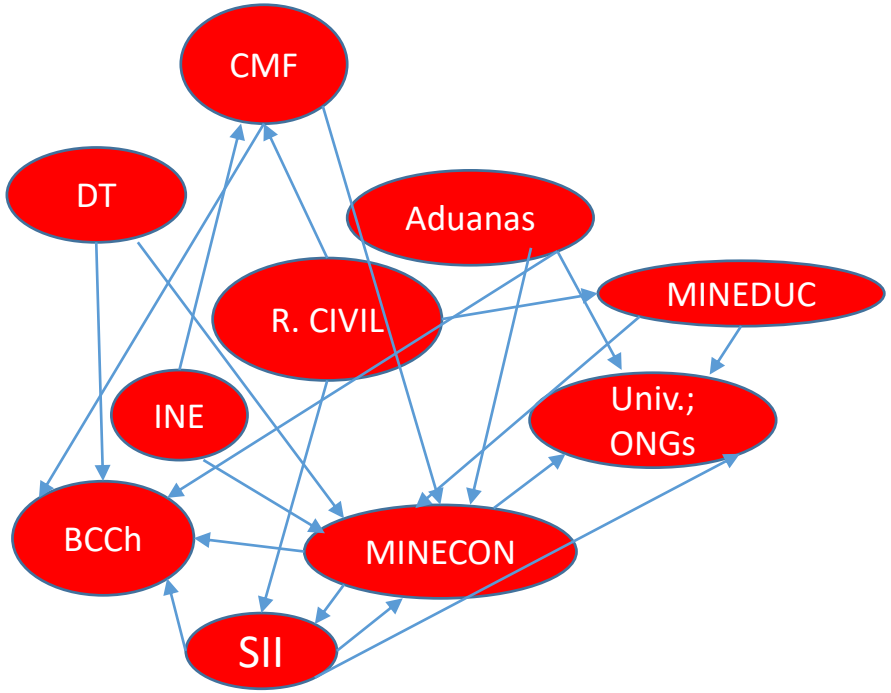
Obstacles

4. **Cooperation and perception**

Real cost of implementing micro-data agenda: limits to cooperation between institutions, due to legal obstacles and/or concerns about the public´s perception

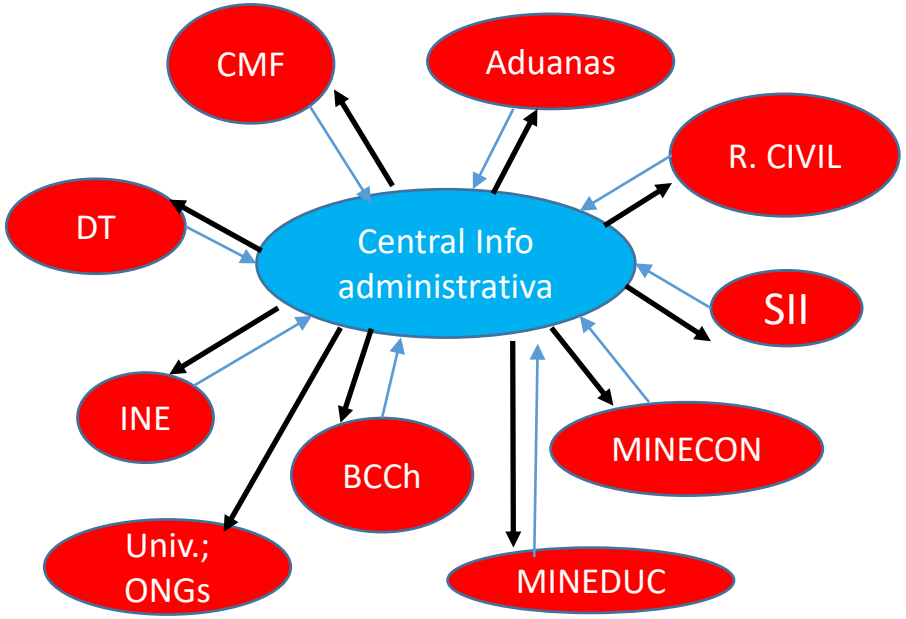
- Most government branches collect data to fulfill individual mandates
  - IRS gathers data with tax-collecting purposes only.
  - INE conducts surveys to update aggregate series (demographics; labor market; activity).
  - Public adm. branches gather info related to specific objectives (M. labor: compliance with labor code; M. education: monitoring schooling outcomes; N. Registry: update demographic info).
    - **--> Little interest in merging info to understand deeper economic interactions.**
- Legal constraints (and their “interpretations”) give ample room for obstructing cooperation
  - Data protection laws impede data merging (or create enough ambiguity to justify inaction).
    - Statistical secrecy; tax-info secrecy; banking-info secrecy.
- In most countries, the public is suspicious of individual data use by govt. authorities (e.g., 1984)
  - In our case, recent progress is largely due to the excellent disposition of the IRS to share data with the CBC, *under highest standards of anonymity and data protection*.
  - But, how can this model be replicated or improved elsewhere?

Optimal model: information flows towards the center. Challenge: convince the public that this is more efficient, safer, and significantly improves quality of govt. services and crisis response



Decentralized model: Chile (ex - CBC) & most countries

- Inefficient & expensive: N*(N-1) potential bilateral agreements to sign.
- Information is partial and incomplete at each node.
- Unsafe: multiple data sharing protocols, ad-hoc rules, increases scope for data filtrations/misuse.



Centralized model: Denmark

- Efficient: N relationships, all facing towards one data aggregator.
- Information is complete at central node. Can then be accessed by all participants according to their needs/mandates.
- Safe: each institution shares information with the central counterpart ONLY, under common standards and safety protocols.

Thanks!