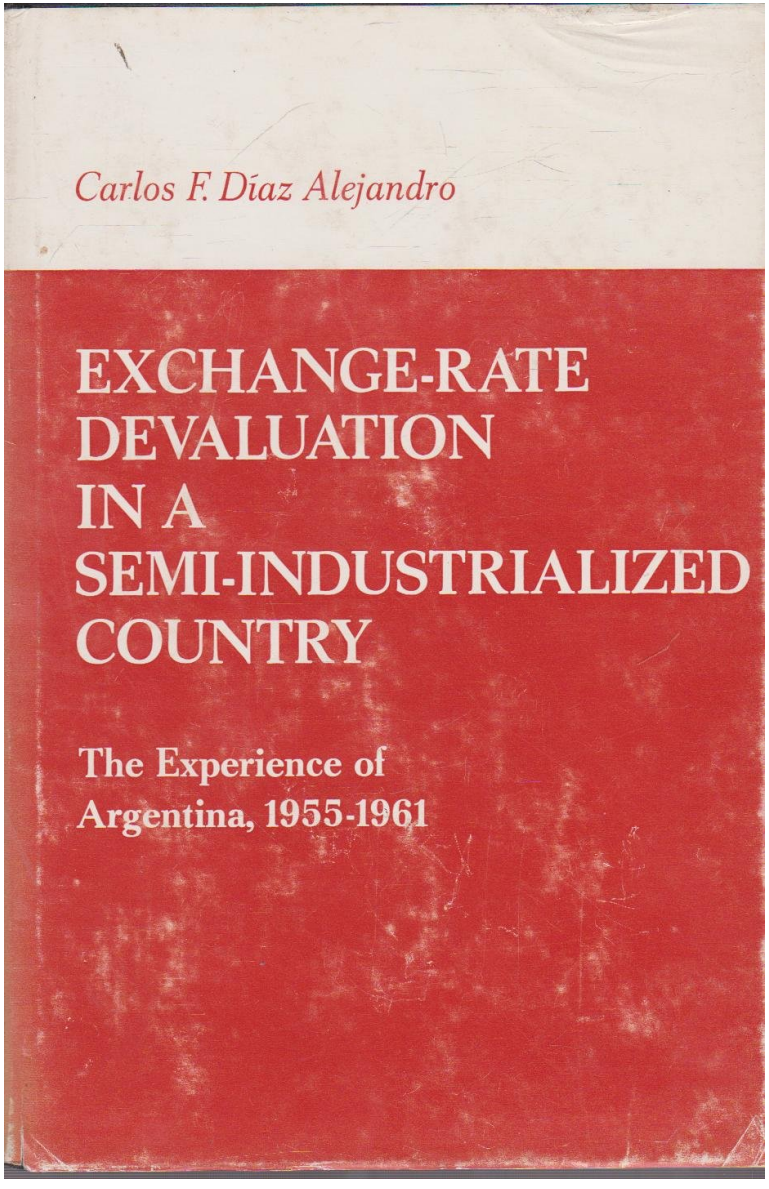
A Reconstruction of Argentina's Consumer Price Index using a Dynamic Factor Model

Rodolfo G. Campos

Banco de España

Presentation at BCRA
Buenos Aires, November 24, 2017

Motivation



Carlos F. Díaz Alejandro
(1937–1985)

- In even years, a prize in his name is given out at LACEA

Carlos F. Díaz-Alejandro: "Elegí Argentina porque tenía mejores estadísticas"

Cuando estaba terminando sus estudios en el MIT, buscando tema de tesis bajo la tutela de Charles Kindleberger, se interesó en encontrarle una explicación al hecho de que, contrariamente a lo que entonces esperaban los economistas, luego de algunas devaluaciones se habían producido recesiones y no aumentos en el producto bruto. Tuvo ante sí las alternativas de estudiar los casos de México y Argentina. "Elegí Argentina porque tenía mejores estadísticas".

de Pablo, Juan Carlos (1995): *"Apuntes a mitad de camino"*, Ed. Macchi, Ch. 14.

Objective

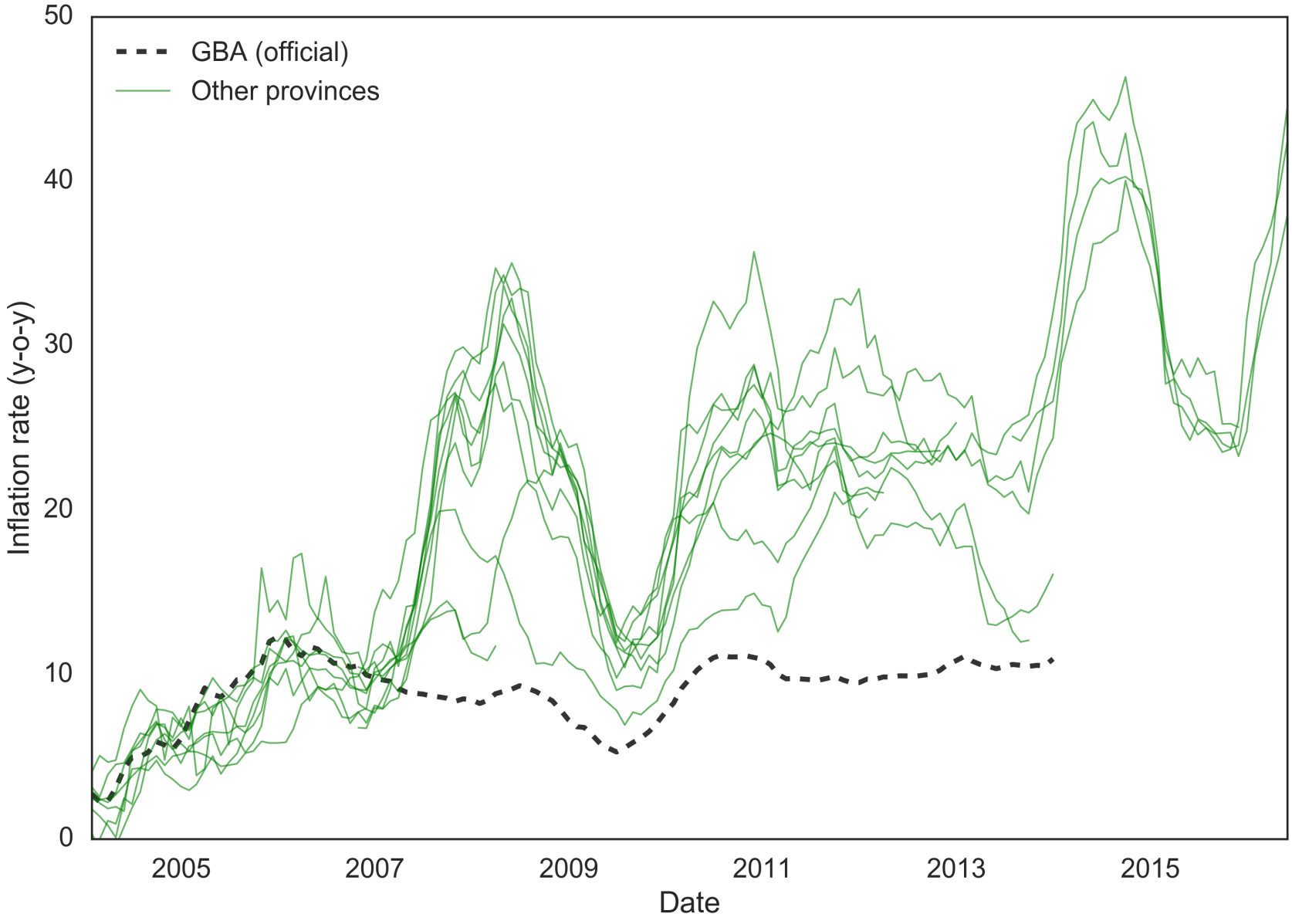
- Argentina's main inflation series (IPC-GBA) started to be unreliable in Jan 2007
- The objective of the paper is to **reconstruct IPC-GBA for the period Jan 2007 – May 2016 using the information contained in provincial inflation rates**.
- How?
  1. Treat unreliable IPC-GBA data as missing observations
  2. Use the techniques of Mariano and Murasawa (2003) to estimate a dynamic factor model with missing observations
  3. Nowcast IPC-GBA for the period for which observations are missing

The Evidence

Table: Monthly inflation rates before and after the intervention

	Period		Average		Std Dev	
	Start	End	Before	After	Before	After
GBA	2003-01	2013-12	0.63	*0.75*	0.37	*0.19*
Ciudad BA	2012-07	2016-05	—	2.45	—	0.99
Chubut	2003-01	2012-10	0.51	1.69	0.64	0.59
Jujuy	2003-01	2012-01	0.53	1.77	0.65	1.07
La Pampa	2003-01	2013-09	0.56	1.41	0.52	0.74
Mendoza	2003-01	2012-03	0.59	1.15	0.50	0.57
Misiones	2003-01	2012-12	0.59	1.70	0.59	0.89
Neuquén	2003-01	2016-05	0.58	2.06	0.74	0.96
San Luis	2005-10	2016-05	0.72	2.01	0.65	1.04
Santa Fe	2005-10	2013-12	0.88	1.42	0.60	0.65
Tierra del Fuego	2003-01	2015-11	0.67	1.89	1.24	0.83
Tucumán	2003-01	2008-03	0.58	0.93	0.54	0.38
Average w/o GBA			0.62	1.68	0.67	0.79
Composite			0.58	1.89	0.42	0.82

Figure: Inflation rates in Argentine provinces



Timeline

1/1934 IPC-GBA is started.

1/2007 Bevacqua, director of the Prices Department, is suspended.

2/2007 Government intervenes INDEC. **First manipulated IPC-GBA is published**. Bevacqua is fired.

3/2007 Director of INDEC resigns.

3/2011 Private consultants are fined for producing their own CPIs.

2/2012 *The Economist* stops reporting the IPC-GBA.

7/2012 IPC Ciudad-BA is started by the city of Buenos Aires.

2/2013 IMF issues a "declaration of censure" on Argentina's statistics.

12/2013 **IPC-GBA is stopped**.

1/2014 IPC-Nu is started.

10/2015 IPC-Nu is stopped.

5/2016 **IPC-GBA is restarted**.

11/2016 IMF removes the "declaration of censure" on Argentina's statistics.

Methodology

The (standardized) monthly inflation rate y_{it} of province i in month t can be decomposed into a common factor and an idiosyncratic component:

$$y_{it} = \gamma_i f_t + u_{it} \tag{1}$$

The factor and the idiosyncratic components follow autoregressive processes of orders p_f and p_i:

$$\begin{aligned} u_{it} &= \sum_{s=1}^{p_i} \psi_{is} u_{t-s} + \varepsilon_{it} \\ f_t &= \sum_{s=1}^{p_f} \phi_s f_{t-s} + \eta_t. \end{aligned} \tag{2}$$

State-space representation

Equations (1) and (2) can be expressed in a static (or stacked) state-space representation by a measurement equation:

$$Y_t = H\xi_t + W_t, \tag{M}$$

with $W_t \sim N(0, R)$ and a transition equation:

$$\xi_t = F\xi_{t-1} + V_t, \tag{T}$$

with $V_t \sim N(0, Q)$.

- If the time series in Y_t do not have missing values, then the estimation can be performed via standard maximum likelihood using the Kalman filter
- Notice that in this application H does not depend on t and that W_t is a vector of zeros

Estimation with missing values

- If there are missing values in Y_t, then the estimation can be performed by modifying the measurement equation in the state-space representation as suggested by Mariano and Murasawa (2003)
- Missing observations are replaced by iid random draws $z \sim N(0, \sigma_z^2)$ that are independent of all parameters in the state-space representation
- Because of independence, the likelihood function of the observed data and that of the data with missing values replaced by random draws will be equivalent up to scale

Substitute:

$$Y_{it}^{+} = \begin{cases} Y_{it} & \text{if } Y_{it} \text{ is observable} \\ z_t & \text{otherwise} \end{cases}$$

$$H_{it}^{+} = \begin{cases} H_i & \text{if } Y_{it} \text{ is observable} \\ \mathbf{0} & \text{otherwise} \end{cases}$$

$$W_{it}^{+} = \begin{cases} 0 & \text{if } Y_{it} \text{ is observable} \\ z_t & \text{otherwise} \end{cases}$$

$$R_{iit}^{+} = \begin{cases} 0 & \text{if } Y_{it} \text{ is observable} \\ \sigma_z^2 & \text{otherwise} \end{cases}$$

Replace the measurement equation with

$$Y_t^{+} = H_t^{+}\xi_t + W_t^{+} \qquad \text{(M')}$$

and estimate as usual with Kalman filter and ML

Scaling

- The scale of the common factor is not identified
- All series are standardized before the estimation
- Recall that the mean and standard deviation differ before and after the intervention
- What is the appropriate scale for the factor and the GBA forecast in the post-intervention period?

Table: Monthly inflation rates before and after the intervention

	Period		Average		Std Dev	
	Start	End	Before	After	Before	After
GBA	2003-01	2013-12	0.63	*0.75*	0.37	*0.19*
Ciudad BA	2012-07	2016-05	—	2.45	—	0.99
Chubut	2003-01	2012-10	0.51	1.69	0.64	0.59
Jujuy	2003-01	2012-01	0.53	1.77	0.65	1.07
La Pampa	2003-01	2013-09	0.56	1.41	0.52	0.74
Mendoza	2003-01	2012-03	0.59	1.15	0.50	0.57
Misiones	2003-01	2012-12	0.59	1.70	0.59	0.89
Neuquén	2003-01	2016-05	0.58	2.06	0.74	0.96
San Luis	2005-10	2016-05	0.72	2.01	0.65	1.04
Santa Fe	2005-10	2013-12	0.88	1.42	0.60	0.65
Tierra del Fuego	2003-01	2015-11	0.67	1.89	1.24	0.83
Tucumán	2003-01	2008-03	0.58	0.93	0.54	0.38
Average w/o GBA			0.62	1.68	0.67	0.79
Composite			**0.58**	**1.89**	**0.42**	**0.82**

$$\hat{\pi}_t^{ARG} = \begin{cases} 0.58\% + 0.42\% \times \hat{f}_t^0, & t < T^* \\ 1.89\% + 0.82\% \times \hat{f}_t^0, & t \geq T^* \end{cases}$$

Table: Monthly inflation rates before and after the intervention

	Period		Average		Std Dev	
	Start	End	Before	After	Before	After
GBA	2003-01	2013-12	**0.63**	*0.75*	**0.37**	*0.19*
Ciudad BA	2012-07	2016-05	—	2.45	—	0.99
Chubut	2003-01	2012-10	0.51	1.69	0.64	0.59
Jujuy	2003-01	2012-01	0.53	1.77	0.65	1.07
La Pampa	2003-01	2013-09	0.56	1.41	0.52	0.74
Mendoza	2003-01	2012-03	0.59	1.15	0.50	0.57
Misiones	2003-01	2012-12	0.59	1.70	0.59	0.89
Neuquén	2003-01	2016-05	0.58	2.06	0.74	0.96
San Luis	2005-10	2016-05	0.72	2.01	0.65	1.04
Santa Fe	2005-10	2013-12	0.88	1.42	0.60	0.65
Tierra del Fuego	2003-01	2015-11	0.67	1.89	1.24	0.83
Tucumán	2003-01	2008-03	0.58	0.93	0.54	0.38
Average w/o GBA			0.62	1.68	0.67	0.79
Composite			**0.58**	**1.89**	**0.42**	**0.82**

$$\hat{\pi}_t^{GBA} = \begin{cases} 0.63\% + 0.37\% \times \hat{y}_t^{GBA}, & t < T^* \\ 0.63\% \times \frac{1.89}{0.58} + 0.37\% \times \frac{0.82}{0.42} \times \hat{y}_t^{GBA}, & t \geq T^*, \end{cases}$$

Miscellaneous details

- I use monthly inflation rates as observations
- I seasonally adjust monthly inflation rates using X13-ARIMA
- I restrict the sample to 01/2013 – 05/2016
- IPC-GBA set to missing starting in 01/2007
- I used the Internet Archive (`https://archive.org/web`) to download provincial CPIs whenever they were erased ex-post

Results

Province	Factor Loading	Asymptotic S.E.
GBA	0.401	(0.200)
Ciudad BA	0.385	(0.065)
Chubut	0.476	(0.055)
Jujuy	0.449	(0.053)
La Pampa	0.394	(0.056)
Mendoza	0.548	(0.076)
Misiones	0.478	(0.050)
Neuquén	0.443	(0.041)
San Luis	0.678	(0.060)
Santa Fe	0.665	(0.105)
Tierra del Fuego	0.368	(0.042)
Tucumán	0.518	(0.108)

Table: Estimated AR coefficients of the idiosyncratic component u_{it}

Province	ψ_{i1}	Asymptotic S.E.	ψ_{i2}	Asymptotic S.E.
GBA	0.678	(0.189)	0.067	(0.158)
Ciudad BA	0.379	(0.144)	0.516	(0.156)
Chubut	0.131	(0.127)	0.002	(0.122)
Jujuy	0.128	(0.106)	0.099	(0.102)
La Pampa	0.263	(0.098)	-0.038	(0.092)
Mendoza	0.435	(0.119)	0.206	(0.112)
Misiones	-0.038	(0.105)	0.055	(0.101)
Neuquén	-0.180	(0.102)	-0.006	(0.102)
San Luis	0.477	(0.118)	0.402	(0.124)
Santa Fe	0.475	(0.122)	0.253	(0.109)
Tierra del Fuego	-0.150	(0.087)	-0.028	(0.086)
Tucumán	0.466	(0.138)	0.348	(0.147)

Table: Estimated variance of the error term ε_{it} in the idiosyncratic component

Province	σ_i^2	Asymptotic S.E.
GBA	0.497	(0.114)
Ciudad BA	0.227	(0.056)
Chubut	0.265	(0.040)
Jujuy	0.344	(0.051)
La Pampa	0.464	(0.061)
Mendoza	0.271	(0.044)
Misiones	0.314	(0.046)
Neuquén	0.318	(0.043)
San Luis	0.168	(0.042)
Santa Fe	0.401	(0.069)
Tierra del Fuego	0.586	(0.070)
Tucumán	0.335	(0.073)

Figure: The common factor

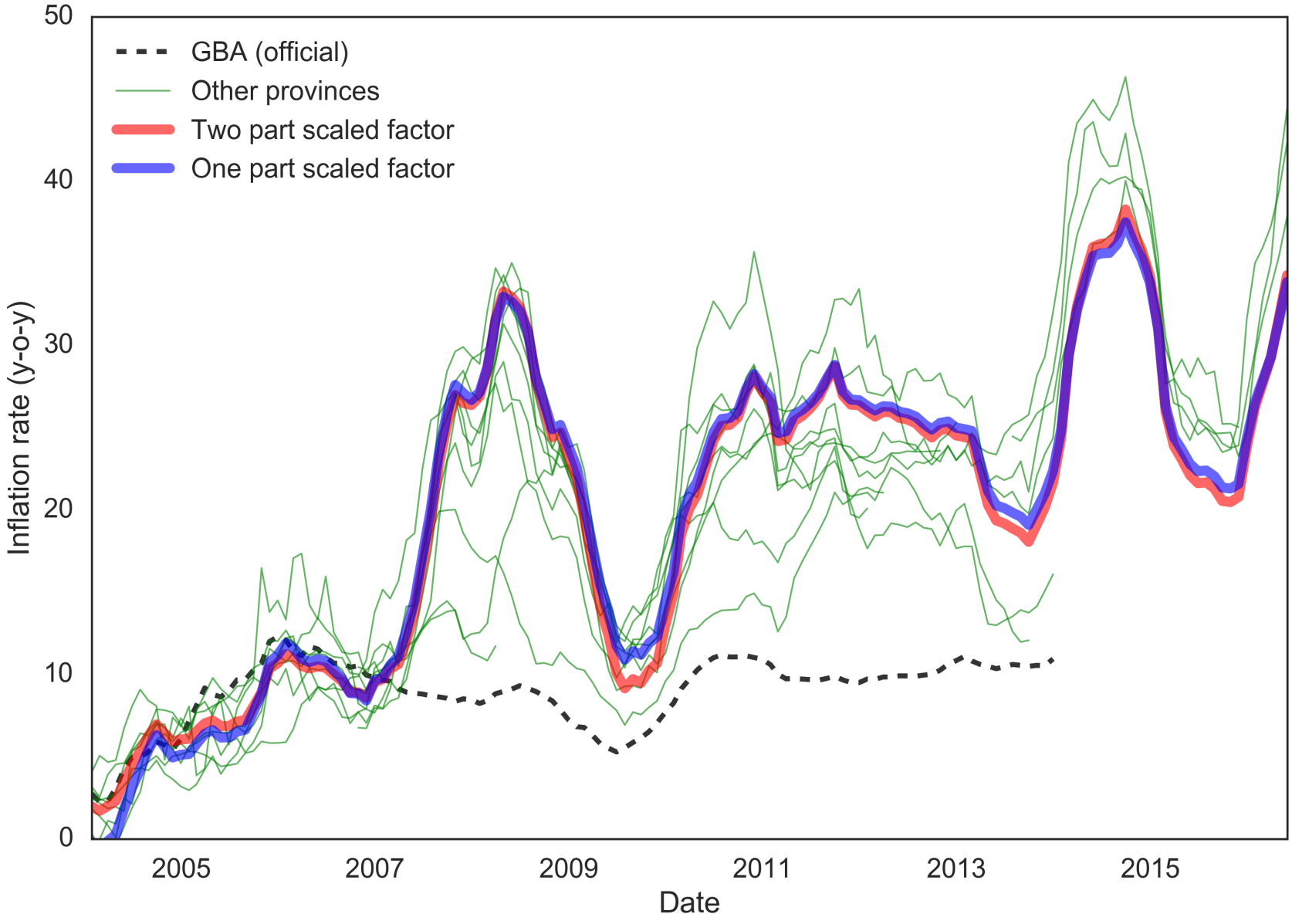


Figure: Estimated inflation for GBA

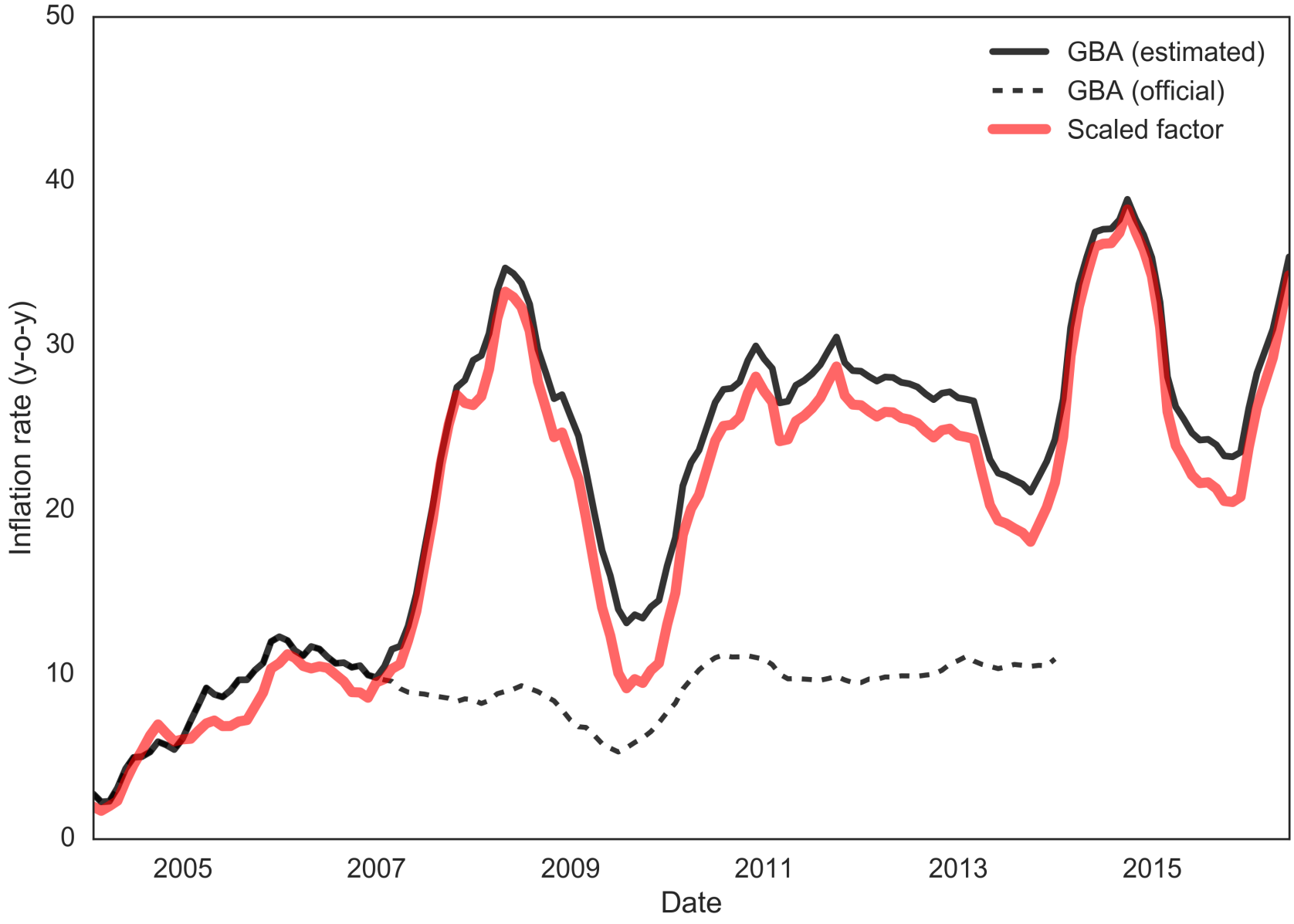
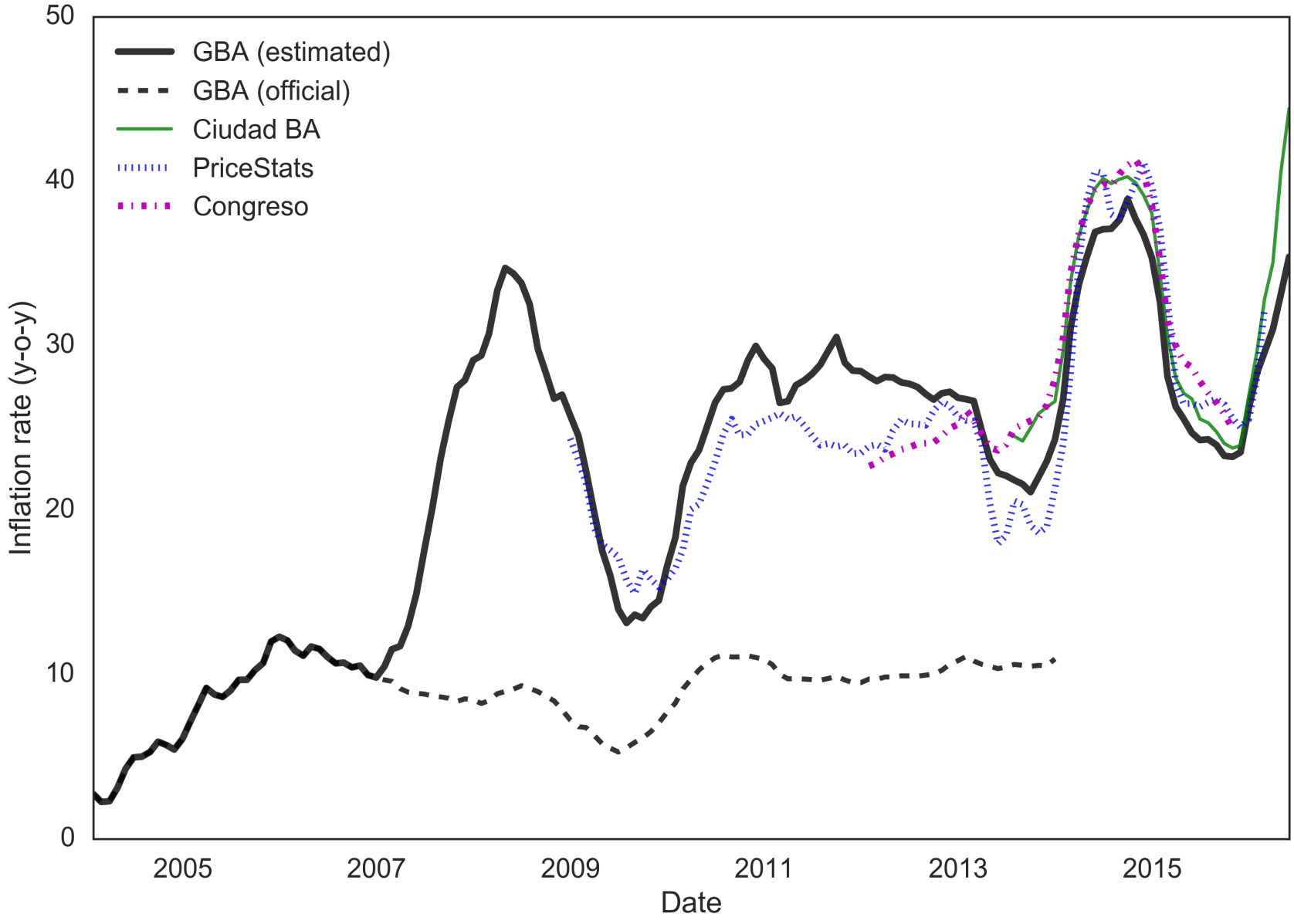


Figure: Comparison with alternative inflation measures



Some comparisons with the official CPI

- As expected, the reconstructed CPI for GBA delivers inflation estimates that exceed official inflation rates after 2007
- Relative to December 2006, the official CPI implies an accumulated rise in prices of **8%** by December 2007 and **87%** by December 2013, when it was discontinued
- In comparison, I find that prices had risen **29%** by December 2007 and **295%** by December 2013
- Prices also kept rising strongly after the official index was discontinued. Over the whole period that goes from December 2006 until May 2016, I find that prices rose by **880%**

Conclusion

- I estimated the CPI for GBA for the period Jan 2007 – May 2016 extracting the latent common factor in provincial inflation rates
- This index coincides with the official CPI before Jan 2007 (by design)
- Thanks to this index there is **a continuous CPI from 1934 to the present that can be used to transform any Argentine time series from nominal into real terms**

Robustness check: using only the 3 longest CPIs

Figure: Comparison using only Neuquén, San Luis, Tierra del Fuego