UNIVERSIDAD TORCUATO DI TELLA
Escuela de Gobierno

ESCUELA DE GOBIERNO

DOCUMENTOS DE TRABAJO 2021/03

# A NEW LABOR CONFLICT INDEX FOR ARGENTINA: PRELIMINARY FINDINGS

DARÍO JUDZIK
EDUARDO LEVY YEYATI
MARTÍN MONTANÉ

FEBRERO 2021

Documentos de trabajo: https://bit.ly/2REorES

UTDT: Av. Figueroa Alcorta 7350, C1428BCW Buenos Aires, Argentina

# A new Labor Conflict Index for Argentina: Preliminary Findings

Darío Judzik (CEPE, School of Government, UTDT) [1]
Eduardo Levy-Yeyati (School of Government, UTDT)
Martín Montané (CEPE, School of Government, UTDT)

February 2021

**Abstract**

Labor conflict has been broadly neglected by economists as a sociological issue with limited impact. In this paper, we present a new Labor Conflict Index based on the application of text mining techniques to media data, document that the results are consistent with the existing official data on labor conflicts, and report our preliminary findings. Labor conflict correlates positively (with a 3-quarter lag) with the business cycle, and it precedes by 2 quarters increases in real wages, indicating that conflict may play a role in the procyclical catch up of real wages during expansions.

[1] Corresponding Author. CEPE, School of Government, Universidad Torcuato Di Tella, Argentina. E-mail: djudzik@utdt.edu.

## 1. Labor conflict and economic performance

Labor conflict has been broadly neglected by economists as a sociological issue with limited impact. In this paper, we argue that the incidence of labor conflict in its diverse forms is systematically associated with the evolution of labor market and macroeconomic variables such as production, wages and employment. In particular, we find that conflict correlates positively with economic activity (with a 3-quarter lag) and that it precedes increases in real wages (by 2 quarters), indicating that it may play an amplifying role in the procyclical behavior of real wages.

In order to explore the labor market and macroeconomic impact of labor conflict we construct a labor conflict index (LCI) for Argentina applying machine learning techniques to media reports, and analyze the empirical links between the behavior of the LCI, the business cycle, and labor related outcomes. First, we compare our index with the available official data on labor conflict and find a robust positive correlation. The LCI benefits from being a considerably larger time series and from being a summary index of the incidence of labor conflict (as opposed to official data that comprise several partial indicators). Next, we explore correlation of the index with the business cycle, with wages and with labor demand expectations. We find the LCI to be procyclical (it correlates positively with the level of economic activity measured as the cyclical component of GDP) and to influence wages (a higher LCI is significantly correlated with a higher real wage a few months down the road), indicating that conflict tends to pay off in terms of labor income. Finally, we find that the spread of conflict increases with higher labor demand expectations (as captured in firm surveys).

It is important to remark that the LCI applies almost exclusively to the private formal sector. We are not taking into account independent or informal labor sectors. As a dual labor market with considerable informality, insiders-outsiders[2] considerations apply to the wage bargaining scheme and labor union representation. Another aspect to keep in mind is that there is a gap between workers actually involved in labor conflict and workers covered or benefited by union action (being the latter a much larger group).

[2] See Lindbeck and Snower (2001).

There is a long lasting debate regarding the economic influence of trade unions. They impose a cost on the production process (e.g. higher wages and restrictions on dismissals) but the controversy on the size of this cost is still ongoing (DiNardo and Lee, 2004). From a policy perspective, the relevant question is if this cost is small in relation to the socioeconomic benefits of labor protection, or is it too high to bear in proportion to its benefits in terms of aggregate welfare. Of course this issue cannot be tackled with a binary yes/no approach. That said, we decidedly need further evidence on how trade union action and labor conflictivity actually affect economic outcomes.

Empirically, the existing evidence on the role of unions and labor conflict on macroeconomic outcomes is mixed. An important bulk of research is summarized in a meta-analysis by Doucouliaglos et al. (2018). They find that the aggregation of results shows that trade unions have no effect on productivity in manufacturing but may have a significant positive effect on industries like construction. They also find that unions reduce investment in physical and intangible capital and argue that the sector-level (and overall small) productivity growth associated with unions does not compensate for union wage increases. Still individual empirical studies offer heterogeneous results and a range of possible conclusions.

In line with this heterogeneity, researchers have found that the duration or the incidence of labor conflict may be weakened or broadened during economic expansions, depending on specific circumstances. Devereux and Hart (2011) show that one can find an analytical rationale to support both a procyclical and a countercyclical pattern. A plausible explanation for this ambiguity points at the fact that the duration of conflictivity may be countercyclical while its incidence may be procyclical (hence, the importance of the selection of a proxy for conflict). Devereux and Hart report in their empirical analysis that labor conflict increases with higher economic activity and aggregate demand. This paper contributes to this discussion by providing new evidence on wider labor conflict as the economy enters a growth cycle. We argue that considering a sufficiently large period of time, systematically labor conflict is higher as businesses boom.

The procyclicality of labor conflict has an intuitive rationale: conflict increases with aggregate production because in periods of economic expansion there are higher stakes – i.e., more wealth to distribute– and, as a result, both sides of the distributional struggle

tighten their bids. Moreover, an expansion lowers the cost of unemployment (as unemployment spells get shorter) and rises the opportunity costs of a strike for the firm, thus strengthening the bargaining power of unions, which suggests that conflict has a decent chance of paying off in terms of wage or labor conditions in times of bonanza. But one could build a similarly convincing argument for countercyclicality by arguing that a recession often pushes workers closer to their reservation threshold (for example, after real wage or payroll adjustments) and that businesses find it less costly to stop production and face a strike (making conflicts more likely than a smooth negotiation). Therefore, the importance of empirical evidence to point policymakers decisively in the right direction. Labor policy, mainly labor relations regulation, can be improved in its goal of achieving higher aggregate welfare by an improved understanding of the relation between labor conflict and economic outcomes.

The connection between labor unrest and its returns in terms of working conditions is not the same across time or industries (Silver, 2003), but also it is predominantly asymmetric in dual labor markets. When there is a clear divide between at least two groups of workers, and one of them is protected by labor law and unions, and the other one is not (or has much lower protection), then the economic effects due to the perception of labor conflictivity affects only one of those groups.

Argentina as a dual labor market is a relevant case study for the international debate on labor issues regarding its tensions between two sectors: on the one hand traditional salaried and protected employment, and on the other hand autonomous self-employment, which gathers liberal high-skilled professionals with informal low skilled workers. Conflictivity is usually conceived as labor conditions bargaining leading up to strikes or other forms of protest. In countries with dual labor markets, this represents only one part of the labor force. For instance, Argentina´s labor conflict index makes sense almost exclusively for formal employment protected by law and union power, roughly half of the labor force in Argentina, with a declining trend. The identification of the segmentation between salaried protected workers and autonomous workers is crucial to see that the distributive struggle held by unions affects only one (and decreasing in size) side of that cleavage.

In this study, we contribute to this debate in two ways. First, we elaborate a methodology for the construction of a labor conflict index (LCI) and apply it to Argentina by text mining news from the main nationwide media sources. The index proxies labor conflict incidence by reporting, at monthly frequencies, the proportion of total uploaded news that include mentions to labor issues and conflicts. Next, we correlate the index with economic activity and the evolution of wages. We do so on monthly and quarterly frequencies, and for the whole sample (1996-2020) as well as for the subperiod with active collective bargaining (2004-2020). We find a procyclical pattern for labor conflict, whereby labor conflict in the form of strikes and other protests increases in the higher part of the business cycle.[3] In addition, we find that, on average, conflict pays off: a higher average real wage follows a few months after a peak in labor conflicts.

Although there may be a differentiation between conflict due to pay (e.g. wage) and non-pay issues (e.g. labor conditions, schedules, and others), in this study we are particularly interested in the connection between aggregate labor-related conflictivity and macroeconomics. Another branch of the labor-relations literature analyzes how the political set up is associated to the dynamics between labor unrest and economic results (e.g. Etchemendy, 2019)[4]. In turn, the questions we address are: what is the relationship between the business cycle and labor conflict? Is there a money reward to labor conflict? How does conflict vary with improving expectations in the labor market? It is important to recall that Argentina has had a significant inflation rate since 2009, and that inflation is both distortive and tightening of the collective bargaining process. In the turmoil of a long lasting recession, and with over 40% inflation rate in the last few years, questions of

3 While, as Franzosi (1995) points out, identifying the direction of causality between social unrest in the form of labor conflict and the macroeconomic juncture is hard, relevant stylized facts can be derived from long enough series displaying significant dynamic correlations as in our case. Naturally, our findings do not imply that there are no conflicts in recessions. We report peaks in the labor conflict index during especially harsh times, such as the 2001-2002 macroeconomic crisis for example. In those specific times, social and political dynamics are key in order to understand the whole scenario.

4 Etchemendy (2019) presents a hypothesis on how some South American labour union representation systems work and how they connect to macroeconomic affairs. Collective bargaining is usually tripartite (between firms, unions and the public administration), and in this scheme unions are as autonomous as the degree of incidence of labor conflict they are capable of generating, even when they are friendly with the incumbent government. So the acceptance of, for example, inflation targeting and other monetary policy goals, is a crucial matter in the degree of conflict one can expect. The same can be said about fiscal policy, government debt and public employment or wage increases. Hence this conflict will be intertwined with monetary and economic consequences.

conflict are relevant for economic policy from labor, monetary and distributional perspectives.

## 2. A labor conflict index

Our contribution is twofold. On the one hand, we construct an index of labor conflict incidence, which becomes useful for future research on the identification of causality in different labor market macroeconomic aspects, such as real wage dynamics, the ongoing inflationary process, and other labor market outcomes including informality and gender gaps.[5]

The construction of the index profits from the tagging of news made by a major national multimedia content and news producer *La Nación*. From 2012 onwards, the paper implemented a tag system that assigns one or more tags to each of the news they post in their online newspaper. Among all the tags, three of them could be proven useful to capturing the labor conflict or bargaining power of the workers: "*Paritarias"*, "*Conflicto gremial"* and *"Paro nacional",* that can be translated respectively as "working conditions bargaining", "union conflict" and "nation-wide strike". We scraped over 1 million news articles from *La Nación* and optimized the parameters of an eXtreme Gradient Boosting (XGB) model to classify news from *La Nación* in the period 2012-2016 based on the full text of the news pieces.

Our methodology can be summarized as follows. First, we removed classic stop words in Spanish and, more importantly, time-specific relevant politicians and renowned labor union leaders. We think that as they are heavily quoted in news related to labor conflict, it could lead to prediction problems when they stopped being quoted. Second, we constructed a Bag of Words (BoW) matrix, in which each row represents a specific news and the columns are the unique words and the cells represent the frequency that the *j* word (column) had in the *i* news (row). Third, we performed a l1 or "row-wise" normalization of the matrix. It replaces the count of the specific word by the percentage of the total word

[5] Further research will explore the role of conflict and union activity on the evolution of sector-level wages. On the other hand, we use text and data mining as our methodology, a growing but still novel approach in the field of social sciences.

count of the news it represents. For example, if there are 20 words and "*paritarias*" appears 3 times, we transform the actual absolute count into 15%[6]. Finally, we optimized the eXtreme Gradient Boosting's parameters using as the training set all the news between 2012 and 2016 and the 2017-2019 period as the testing set (see Figure 1)7

[Figure 1 about here]

After we optimized the parameters, we classified the news of other major newspapers such as *Telam* and *Pagina|12*. It is important to stress that we proactively gathered news from newspapers and news agencies that have explicitly opposite editorial lines in order to control for any reporting or covering bias that they may have. Figure 2 shows the percentage of news associated with these tags since 2017 for all three newspaper media. Levels differ, but the peak-and-trough patterns are comparable which is the main concern.

[Figure 2 about here]

## 3. Is labor conflict procyclical?

### 3.1. The Argentine Labor Market

Salaried work is by far the frequent labor relationship in Argentina, above the mean for Latin America and the Caribbean (74,3% against 62,5% in 2019, according to the World Bank's *World Development Indicators)* but roughly in line with its relative economic development. That said, the share of salaried workers that are registered and subject to collective bargaining, and are therefore relevant to the LCI –that is, the direct beneficiaries from labor demands– is barely over half of the total work force, and around three quarters of total salaried workers. According to Argentina´s Household Survey for the first three quarters of 2020, out of the 72.5% salaried workers, only 53% (33.3% in the private sector) are registered under labor law and subject to collective bargaining. Remaining workers include those informally employed (19.5%), and the self-employed or independent workers (27.5%), which in turn include a majority of low-skilled informal

[6] We chose not to use the td-idf transformation since it further weights the word weight by the frequency of the word in all the documents. Since our aim is to predict retrospectively we think that there could be changes in the use of language that could bias our predictions.
7 We proceeded in this way because that is exactly what we aim to do: identify the relationship between the variables during a fixed period and try to classify the news in another time span.

workers and a minority of high-skilled liberal professionals (estimated at 20.9% and 6.6% respectively). This composition is critical for our analysis, as labor conflict might be focusing exclusively on registered salaried workers, if not solely on those heavily unionized. Additionally, it is worth mentioning that Argentina has a declining share of salaried workers under collective bargaining. Table 1 presents the main trends in these proportions for Argentina in recent years.

[Table 1 about here]

Official data on labor conflict for Argentina starts in 2006 in the form of three time series: number of conflicts, workers involved, and workdays lost. Reassuringly, our LCI presents a large, positive and significant correlation with workdays lost to conflict and with the amount of workers involved. This is an important robustness check for our LCI produced from data mining (see Table 2). The account for conflicts has a non-significant correlation with the LCI, possibly because it has no ponderation for the size and duration of these conflicts and is therefore a poor measure of incidence. By contrast, the LCI is ten years longer (starts in 1996) and is more parsimonious and informative about the intensity of labor market conflicts.

[Table 2 about here]

3.2. Conflict, growth and wages

Our main preliminary results are presented in Table 3. The purpose of these observations is to characterize the stylized facts of the cyclical patterns, if any, of the long time series (1996 to 2020 both in quarterly and monthly frequency as the LCI, or otherwise as many observations available).

[Table 3 about here]

According to our LCI, labor conflict in Argentina is procyclical: the LCI shows a significant correlation with lagged economic activity. In other words, an increase in economic activity is positively correlated with a higher incidence of labor conflict a few months later, in line with the international evidence reported by Devereux and Hart

(2011). This correlation is higher for the 2004-2020 period, especially for quarterly data (Figure 3). This is due to core economic changes around 2003/2004, including the reactivation of collective bargaining (resumed under new regulations) and a change in the monetary regime with a gradual resurgence of inflation, which became substantial to wage negotiations. Hence, the 2004-2020 period is especially relevant for the macroeconomic dynamic of labor conflict. Plus, the quarterly frequency smooths out some of the noise in monthly time series.

[Figure 3 about here]

The correlation of conflict and real wages peaks after an increase in conflict: we find a positive and significant correlation between the ICL and the mean real wage. This evidence suggests that labor conflicts may originate from wage misalignments and they may eventually pay off for the workers involved –or at least perceived to do so-, since at this stage we cannot identify conflictive activities and see a gain relative to the generally cyclical recovery in real wages.

Interestingly, we also find relevant correlations between the LCI and business expectations for the near future. We use data from the Labor Indicators Survey run monthly by the Ministry of Labor, Employment and Social Security (EIL). In this survey, businesses are asked if they are willing to increase, decrease or keep constant their level of employment. The expectation of hiring correlates positively with the LCI.

It is important to mention that we checked for robustness with different sub-periods samples and with different combinations of news and media agencies as sources for labor news. We find no systematic bias, since levels of conflict coverage may vary, but in first differences, the peaks of labor conflict are very similar (correlation coefficients of between 0.70 and 0.86 for the resulting index using different news sources with opposite editorial lines).

## 5. Concluding remarks

We applied machine learning to the construction of evidence that is consistent with existing data, but longer and more nuanced and, as such, useful for labor market analysis. Our Labor Conflict Index can be replicated for any country or region. Preliminary results from statistically significant dynamic correlations are consistent with a short-run Phillips curve in the sense that labor bargaining power increases in economic expansions and pushes real wages higher. However, a more detailed sector-specific analysis is needed to examine whether labor conflict is ultimately validated with ex-post higher relative wages for the sector. Further research should look in more detail into the channel through which increase in real wages are associated to conflict and whether conflict is used to enhance workers´ bargaining power at the time collective negotiations.

**References**

Devereux, P.J. and R.A. Hart (2011) "A Good Time to Stay Out? Strikes and the Business Cycle", *British Journal of Industrial Relations*, vol. 49, pp. s70-s92.

DiNardo, J. and D. S. Lee, (2004) "Economic Impacts of New Unionization on Private Sector Employers: 1984–2001", *The Quarterly Journal of Economics*, vol. 119(4), pp. 1383-1441.

Doucouliagos H., R. B. Freeman, P. Laroche, and T. D. Stanley. (2018) "How Credible Is Trade Union Research? Forty Years of Evidence on the Monopoly–Voice Trade-Off", *ILR Review*, vol. 71(2), pp. 287-305.

Etchemendy S. (2019) "The Rise of Segmented Neo-Corporatism in South America: Wage Coordination in Argentina and Uruguay (2005-2015)", *Comparative Political Studies*, vol. 52(10), pp. 1427-1465.

Franzosi, R. (1995) *The Puzzle of Strike.: Class and State Strategies in Postwar Italy.* Cambridge University Press, ISBN 0-521-45287-2

Lindbeck, A. and D.J. Snower (2001) "Insiders versus Outsiders", *The Journal of Economic Perspectives*, vol. 15(1), pp. 165-188.

Silver, B. (2003). *Forces of Labor: Workers' Movements and Globalization Since 1870.* Cambridge: Cambridge University Press.

**Tables and Plots**

Table 1.

Main indicators of the job market of Argentina. Source: Permament Household Survey (EPH) of the National Statistics and Census Institute (INDEC)).

| Year | Formal Salaried Worker (private sector) | Informal Salaried Worker (private sector) | **Registered salaried workers (% of total salaried workers in the private sector, %)** | Salaried worker (government or enterprise owned by the government, %) | **Registered salaried workers (% of total salaried workers)** | **Salaried workers (% of total workers )** | Non Salaried Workers | **Share of registered salaried workers (% of total workers)** |
|---|---|---|---|---|---|---|---|---|
| 2003 | 25.8% | 26.8% | **49.0%** | 14.7% | **60.2%** | **67.4%** | 32.6% | **40.6%** |
| 2004 | 27.5% | 27.4% | **50.1%** | 13.9% | **60.2%** | **68.8%** | 31.2% | **41.4%** |
| 2005 | 29.2% | 27.3% | **51.7%** | 14.0% | **61.3%** | **70.5%** | 29.5% | **43.2%** |
| 2006 | 31.3% | 26.7% | **53.9%** | 14.6% | **63.2%** | **72.6%** | 27.4% | **45.8%** |
| 2007 | 34.0% | 25.8% | **56.8%** | 14.5% | **65.3%** | **74.4%** | 25.6% | **48.6%** |
| 2008 | 35.8% | 24.3% | **59.6%** | 14.9% | **67.6%** | **74.9%** | 25.1% | **50.6%** |
| 2009 | 35.7% | 23.5% | **60.3%** | 15.5% | **68.6%** | **74.7%** | 25.3% | **51.2%** |
| 2010 | 36.4% | 22.9% | **61.4%** | 15.7% | **69.5%** | **74.9%** | 25.1% | **52.1%** |
| 2011 | 36.9% | 22.4% | **62.2%** | 15.8% | **70.2%** | **75.1%** | 24.9% | **52.7%** |
| 2012 | 36.4% | 22.7% | **61.6%** | 16.4% | **70.0%** | **75.4%** | 24.6% | **52.8%** |
| 2013 | 36.6% | 21.9% | **62.5%** | 16.2% | **70.7%** | **74.6%** | 25.4% | **52.7%** |
| 2014 | 36.1% | 21.5% | **62.7%** | 16.8% | **71.1%** | **74.4%** | 25.6% | **52.9%** |
| 2015 | 36.4% | 21.2% | **63.1%** | 17.2% | **71.6%** | **74.8%** | 25.2% | **53.6%** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2016 | 35.3% | 21.6% | **62.1%** | 17.0% | **70.8%** | **73.8%** | 26.2% | **52.3%** |
| 2017 | 34.5% | 21.6% | **61.5%** | 17.0% | **70.4%** | **73.1%** | 26.9% | **51.5%** |
| 2018 | 34.9% | 21.2% | **62.2%** | 16.6% | **70.8%** | **72.7%** | 27.3% | **51.5%** |
| 2019 | 34.9% | 21.0% | **62.4%** | 16.0% | **70.8%** | **71.9%** | 28.1% | **50.9%** |
| 2020* | 33.3% | 19.5% | **63.0%** | 19.7% | **73.1%** | **72.5%** | 27.5% | **53.0%** |

* Does not include the fourth quarter.

Table 2: Correlation of the LCI with official labor conflict series

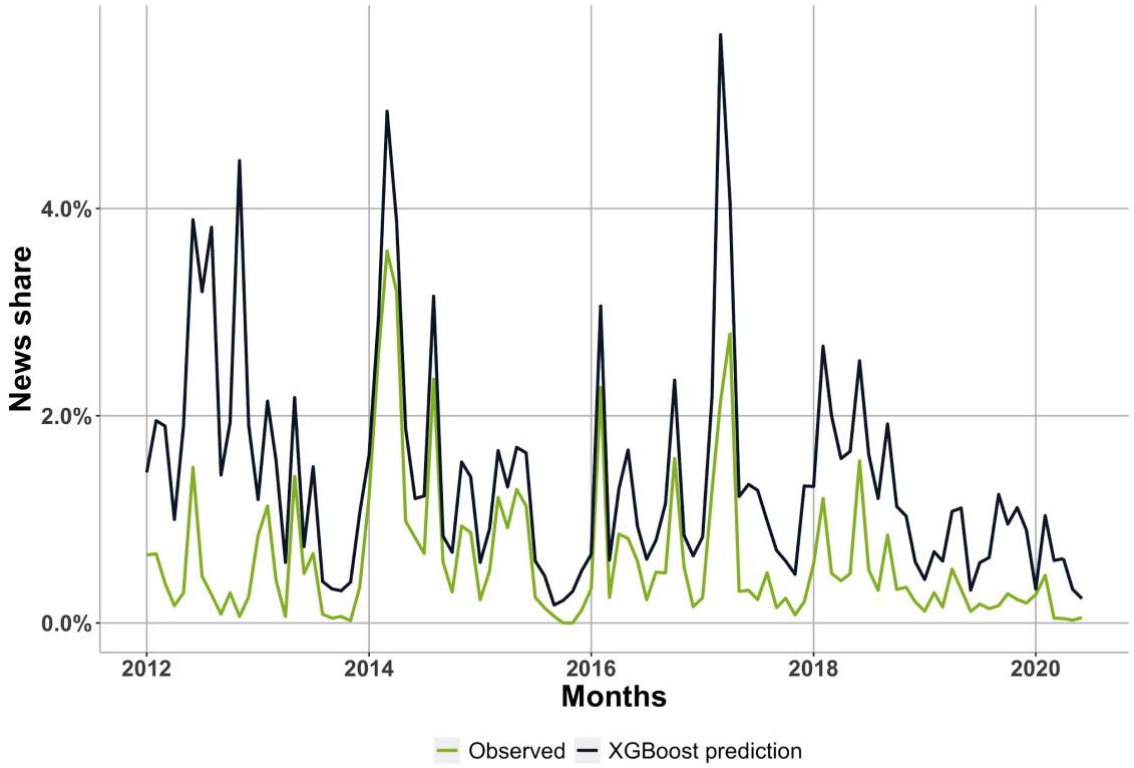
| | | Sample 2004-2020 | | | | | |
|---|---|---|---|---|---|---|---|
| | | Source | Correlation with ICL | Person's p-value | Number of observations | Highest lag/lead in corr. | |
| Quarterly | Strike duration (work days lost) | MTEYSS | 0.44 | 0.00 | 180 | t=0 | |
| | Workers involved in strikes | MTEYSS | 0.49 | 0.00 | 180 | t=0 | |
| Monthly | Strike duration (work days lost) | MTEYSS | 0.46 | 0.00 | 60 | t=0 | |
| | Workers involved in strikes | MTEYSS | 0.45 | 0.00 | | 60 | t=0 |

Table 3: Dynamic correlation of the LCI with economic activity and real wages

| | Sample | | Full sample (1996-2020) | | | | Subsample 2004-2020 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Source | Correlation with ICL | Person's p-value | Number of observations | Highest correl. (t=0, lag, lead) | Correlation with ICL | Test t p-value | Number of observations | Highest lag/lead in corr. |
| Quarterly | Business cycle | INDEC | 0,21 | 0.04 | 96 | lag | 0.33 | 0.01 | 68 | lag |
| | Real wage | SIPA-MTEYSS and combined sources for CPI | 0.07 | 0.48 | 96 | lead | 0.27 | 0.03 | 64 | lead |
| | Expectation of higher labor demand | EIL-MTEYSS | -- | -- | -- | -- | 0.39 | 0,00 | 65 | lag |
| Monthly | Business cycle | INDEC | 0.10 | 0.08 | 292 | lag | 0.14 | 0.05 | 204 | lag |
| | Real wage | SIPA-MTEYSS and combined sources for CPI | 0.03 | 0.64 | 290 | t=0 | 0.23 | 0.00 | 194 | -- |
| | Expectation of higher labor demand | EIL-MTEYSS | -- | -- | -- | -- | 0.23 | 0.00 | 193 | lag |

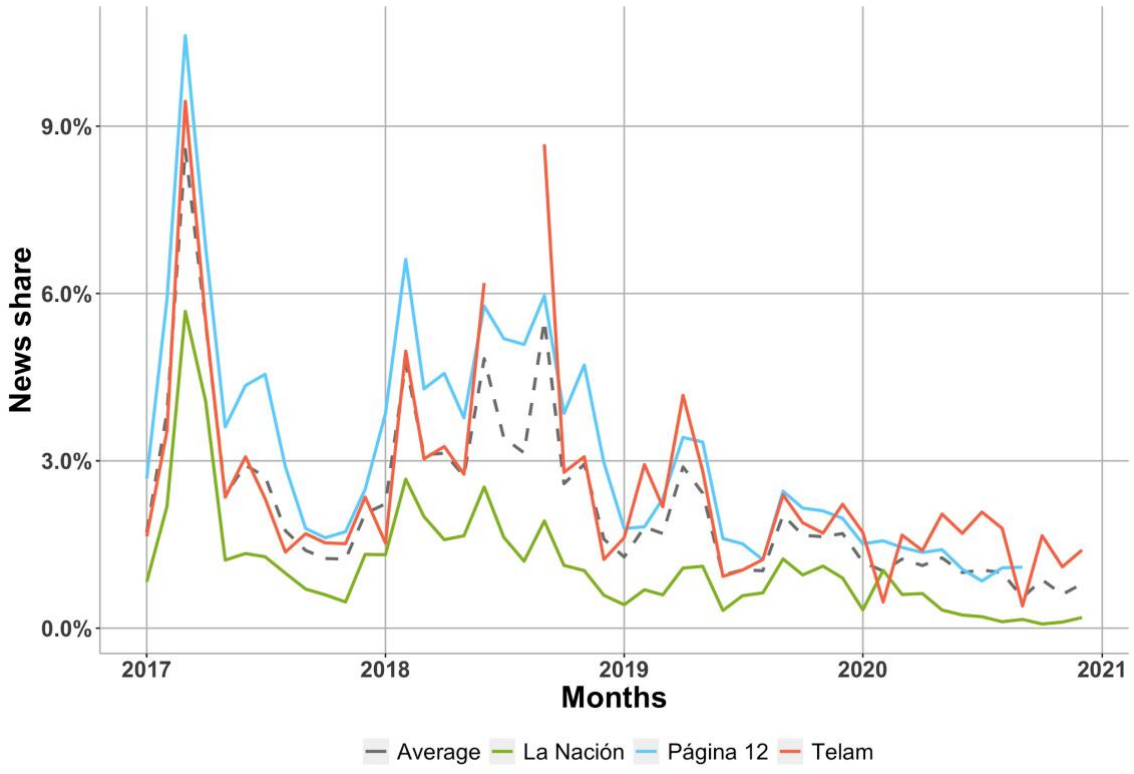
Notes: CPI from alternative sources to INDEC during intervention.

Figure 1



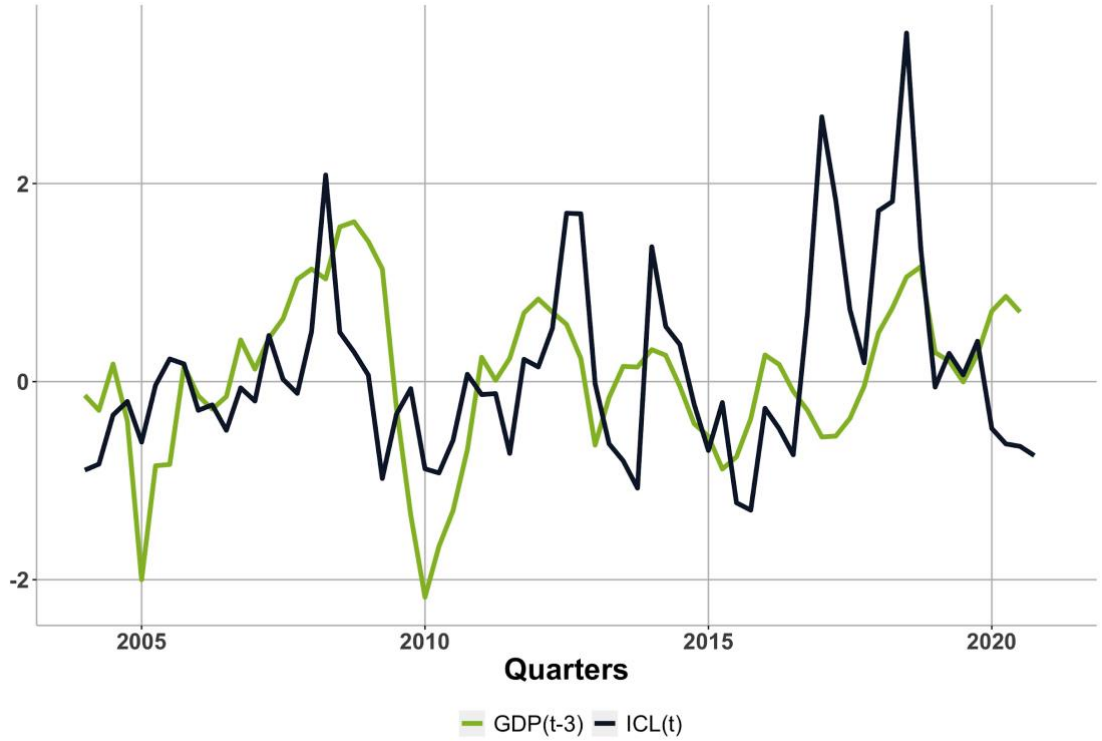

Sources: ICL (CEPE-UTDT).

Figure 2




Source: ICL (CEPE-UTDT).

Figure 3




Sources: ICL (CEPE-UTDT) and INDEC.