Survey Non-response in Covid-19 Times: The Case of the Labour Force Survey

Pierre Brochu* and Jonathan Créchet†

August, 2021

Abstract

With the advent of the Covid-19 pandemic, labour-force survey non-response rates have surged in many countries. We show that in the case of Canada, the bulk of this increase can be explained by the suspension of in-person interviews following the adoption of telework in Federal agencies, including Statistics Canada. Individuals with vulnerabilities to the Covid-19 economic shock—i.e., the young, low-educated, low-salary, low job-tenure individuals, and those working in occupations with low telework potential—have been harder to reach and have been gradually less and less represented in the Canadian Labour Force Survey (LFS) during the pandemic. Using exogenous variation in the assignment of individuals to the different LFS rotations, we present evidence suggesting that the decline in employment and labour-force participation have been underestimated over the March-July 2020 period. We believe, however, that these non-response biases have been moderate when contrasted with the unprecedented severity of the Covid-19 disruption. Furthermore, since attrition only represents a minor part of the non-response increase, we argue that one should not expect additional difficulties when using panels as compared to cross-sectional samples, and when using public-use LFS files instead of restricted-access files. All in all, the LFS remains a reliable data source for analyzing the economic impact of Covid-19 in a timely manner.

*Corresponding author: Pierre Brochu, Department of Economics, University of Ottawa, Ottawa, Canada. Email: pbrochu@uottawa.ca. We are grateful to Louis-Philippe Morin for his invaluable advice and support throughout the project. The results presented in this article reflect our views and interpretations only and not those of Statistics Canada. Our results are based on information that is available in the master files of the LFS, accessible to researchers through Statistics Canada Research Data Centres, and on national LFS non-response rates that are readily available to researchers. All errors and omissions are ours.

†Department of Economics, University of Ottawa. Email: jcrechet@uottawa.ca.

1 Introduction

The propagation of Covid-19 in Canada and throughout the world has resulted in a tremendous and pervasive disruption of economic activity. Statistical agencies have clearly not been spared and their data collection activities have been severely constrained. As a result, survey non-response rates have dramatically surged, raising concerns about the reliability of statistics in Covid times (e.g., Rothbaum & Bee (2021)). Yet, a substantial body of work analyzing the economic impact of Covid-19 has been based on survey data (e.g., Baylis et al. (2020); Forsythe et al. (2020); Lemieux et al. (2020); Jones et al. (2021)). This research has critical public policy implications, making it crucial to evaluate the impact of Covid-19 on this type of data source. In particular, it has been widely documented that the impact of the downturn has been highly heterogeneous across individuals, and as such, a key question is whether non-response is more likely among subgroups with characteristics correlated with Covid-19 labour-market outcomes. This paper presents evidence showing that in the case of the Canadian Labour Force Survey (LFS), non-response has been significantly higher among individuals who have been among the most impacted by the Covid shock.

The LFS has been characterized by persistently high non-response rates since the onset of Covid. Non-response started to rapidly climb as of March 2020 when lockdowns were being imposed, reaching 29.5% in August 2020, and remaining stubbornly high for the remainder of the year. To give some historical perspective, prior to the pandemic the monthly non-response rate had never exceeded 14% since the advent of the modern-day LFS in 1976 (Brochu 2021). As a comparison, the increase in the non-response rate of the U.S. Current Population Survey (CPS) was very sharp but relatively short-lived (see figure 1). Our analysis shows that the major part of the persistent increase in LFS non-response can be attributed to the suspension of in-person interviews at Statistics Canada, as a result of the adoption of telework arrangements in federal agencies. It has made it more difficult for LFS surveyors to contact the most vulnerable individuals over the course of the pandemic.[1] Our results suggest that the youngest, low-educated, low job-tenure, low-salary individuals and those working in occupations with low potential for telework have been the hardest to reach.

Our analysis is based on the master (confidential-use) files of the LFS which provide key information about non-response and more generally, about the information collecting process

[1] The CPS also suspended in-person interviews as of March 2020. However, contrary to the LFS, the CPS restarted in-person interviews in some areas of the country as of July 2020, expanding to the whole country beginning in September 2020. As of January 2020, the LFS had yet to resume in-person interviews. It should be noted that before the pandemic, only a minority of interviews were carried out in person, and this holds true for both LFS and CPS. Finally, it is important to recognize that although the CPS non-response rates did come back down after their initial surge, they did not return to pre-pandemic levels.

for the different individuals appearing in the survey. Specifically, the master data allows for precise tracking of each individual's LFS history, including their assigned rotation, and whether or not they have been reached across months.[2] These features are key to our analysis. For instance, they allow us to show that most of the increase in non-response can be accounted for by *birth* non-response (i.e., the failure to establish initial contact with individuals in a sampled household) rather than *subsequent* non-response (the failure to collect current information on individuals who have already been reached in a prior month).

The information about the type of non-response allows us to trace back the increase in non-response to the suspension of in-person interviews, which historically have played a key role in establishing contact with households. Furthermore, the importance of birth relative to subsequent non-response has implications for the impact of Covid-19. Indeed, subsequent non-response allows for the imputation of labour-market information to individuals based on their pre-existing information, as opposed to birth non-response for which no information is known. Thus, birth non-response, if non-random, induces a change in the structure of LFS samples. For instance, those most mobile might be harder to reach, which might explain why youths and low–job-tenure (conditional on age) individuals tend to be more likely to be interviewed in person (in normal times).

Due to the changes in sample structure resulting from the increase in non-response, we might expect that labour-market estimates have been affected by a non-response bias.[3] To test for the presence of selection based on characteristics correlated with Covid-19 labour-market outcomes, we again make use of the individual's rotation information that is available in our LFS files. Essentially, we compare the impact of Covid on employment and participation across two groups of individuals: those who have been assigned to rotations that entered the LFS just before the onset of Covid-19 (i.e. before March 2020) and those assigned to rotations that joined afterwards. These two groups have been selected according to two different rules: for the former group, in-person interviews were available, but for the latter

[2]As with many labour force surveys, e.g., the CPS (U.S.) and Enquête Emploi (France), the LFS has a rotating sample design. LFS Households are surveyed for six consecutive months, with one-sixth of the households being replaced every month. Said differently, households are assigned to rotations that enter the LFS in a staggered manner. For example, rotations 1 and 2 enter in January and February of each year, only to be replaced by two new groups of households in July and August, respectively. Note that due to household non-response, a household may not be observed in all six months of its rotation's stay in the LFS.

[3]As discussed later in the paper, the LFS weights mitigate but do not completely eliminate, the potential issues generated by the changing sample structure. The LFS weights are adjusted to correct for relative differences in the prevalence of non-response across groups based on a list of key socio-demographic and labour-market variables (including, for example, age, employment, and industries). As such, the weights do not correct for dimensions not taken into account in this adjustment process (e.g., education, occupations, and job tenure).

group, they were not. Assuming that the assignment to rotations is uncorrelated with labour-market outcomes in a given survey month,[4] a different impact across the two groups should be due to differences in structure—a result of their different selection rules.

We find that for the early Covid months, the employment and participation impacts are different across rotations that entered the LFS pre- and post-Covid.[5] Hence, we cannot reject the null hypothesis of an absence of selection along individual characteristics correlated with Covid outcomes. Our (weighted) regressions suggest that estimates based on the entire LFS sample have underestimated the negative impact of Covid: being assigned to a post-Covid rotation is associated with a lower Covid impact, consistent with our evidence that the most vulnerable have been the hardest to reach. The magnitude of the discrepancy is relatively high for the 50-64 year-olds, which suggests that employment and participation losses have been especially underestimated among this group.

A limitation of our approach is that since the group of individuals assigned to pre-Covid rotations is also affected by an increase in non-response, we cannot make conclusions on the size of the bias. Under mild assumptions on the effect of non-responses samples, we can only conclude about the sign of the bias. We believe, however, that biases have moderate magnitude and do not change the narrative for the Covid-19 aggregate labour-market dynamics that has been proposed so far. One of the main reasons to be optimistic is that LFS non-response rates have remained at moderate levels despite a substantial increase: a rate of 30%, as observed in the second half of 2020, would be considered more than acceptable for the greater majority of surveys—even before Covid.[6] However, our results shed light on the importance of being careful when analyzing Covid outcomes of specific subgroups—e.g., the oldest or those belonging to incoming rotations.

Throughout our analysis, we discuss the implications of the pandemic disruption for research based on LFS data. The surge in non-response raises two important questions: first, should researchers expect additional difficulties when carrying out a longitudinal analysis? Building panels requires that one be able to follow individuals over time and, therefore, it imposes additional sample restrictions. However, since subsequent non-response, and, as such, panel attrition is a mild issue (at least when compared to birth non-response) there are no sub-

[4]Given the design of the LFS, there is no reason to believe that as of a particular month, say March, that households assigned to different rotations would have systematically different labour market outcomes.

[5]For ease of notation, we will refer to a pre-Covid rotation as one that joined the LFS prior to the advent of Covid. Similarly, a post-Covid rotation is defined as one that entered the LFS in the Covid-era.

[6]It has been widely documented that non-response has been increasing over the past few decades in Canada and abroad (e.g., Barrett et al. (2014); Meyer et al. (2015); Bernhardt et al. (2021)). Although one did see a modest rise in LFS non-response rates in the 2000s, they stabilized in the 10-12% range over the past decade prior to the Covid outbreak.

stantial differences between panels and cross-sectional samples in terms of their structure, including socio-demographic characteristics. The second question is to what extent it is critical to use the LFS master files (as opposed to public-use files) when focussing on the pandemic period? The master files provide information on data imputation that is not made available in the public-use files. Given the surge in non-response, it is natural to expect that such information is beneficial for producing a robust statistical analysis, but the benefit must be balanced with the administrative costs of accessing master files.[7] And once again, the fact that subsequent non-response is a mild issue implies that the benefit of using master files is not as high as one could expect at first sight regarding data-imputation concerns.

A vast literature has analyzed the economic impact of Covid-19, but much less is known about the impact upon data quality. Information is becoming available to researchers on the impact upon U.S. datasets, whether it be from the U.S. Bureau of Labour Statistics (e.g., U.S. Bureau of Labor Statistics (2020)), third-party data providers (e.g., IPUMS-CPS), or other researchers themselves (e.g., Montenovo et al. (2020); Rothbaum & Bee (2021)). However, little (if any) is known about how Canadian data were impacted.[8] Yet, the LFS has been widely relied upon due to its timeliness (the microdata is published less than a month after it has been collected), and its high frequency (i.e. monthly). It is a workhorse of the Canadian Covid literature,[9] and more generally, of the Canadian labour-market literature.[10] Moreover, researchers who rely on the LFS often use multiple decades of data (e.g., Brochu & Green (2013); Jones & Riddell (2019)). Therefore, understanding how the data were affected by Covid-19, even if the impact turns out to be only short-term in nature, will be of relevance to researchers and policy-makers for years to come. Our paper takes a step in this direction.

[7] When choosing between public-use and master files of the LFS, researchers face a trade-off between easy access and information richness, as discussed in Brochu (2021). Assessing this trade-off is crucial in the case of Covid-time data given that the master files allow for the identification of individuals for whom labour-market information has been imputed due to subsequent non-response (i.e. for whom previous information has been collected and is used for imputing missing information—see the main text for details).

[8] Brochu (2021) provides a very brief overview of how Covid-19 affected non-response in the early stages of the pandemic. It cites an earlier working paper version of this paper, i.e. Brochu et al. (2020), as evidence suggesting that there is reason to be optimistic about the reliability of LFS data during Covid times.

[9] See for example, Achou et al. (2020); Baylis et al. (2020); Béland et al. (2020); Béland et al. (2020b); Blit (2020); Bryan et al. (2021); Gallacher & Hossain (2020); Jones et al. (2020); Koebel & Pohler (2020); Lemieux et al. (2020); Messacar et al. (2021); Qian & Fuller (2020).

[10] See for example, Kroft et al. (2019). Of note, the LFS has been used to examine the impact of a variety of labour market policies such as disability insurance (e.g., Milligan & Schirle (2019)), employment insurance (e.g., Lluis & McCall (2019)) and universal child care benefits (e.g., Schirle (2015)), to name a few.

2 Covid-19 and non-response in the LFS

The main events that marked the propagation of Covid-19 in Canada (and worldwide) are now well known. The first Canadian case was reported at the end of January 2020, and by mid- to late March, provincial governments had shut down all non-essential activities. These restrictions were eased during the summer of 2020, but were partially reimposed amid the second and third waves of the outbreak, in the fall of 2020 and the spring of 2021, respectively.[11] As of April 2021, the country had more than a million confirmed cases, with an associated 23,000 deaths.

The shutdown policies of early spring 2020 also impacted the data collection activities of Statistics Canada. The first significant manifestation of this disruption is the surge in household non-response rates that followed the outbreak. Between February and March 2020, the LFS non-response rate almost doubled, rising from 11.9% to 22.1%. Although there have always been regional differences, i.e. non-response rates tend to be slightly higher in Western Canada, the impact of the pandemic was felt country-wide. This upward trend continued thereafter, reaching 30.7% in September 2020. The impact has been strikingly persistent, despite the successive easing and tightening phases of social-distancing restrictions. As of the beginning of 2021, the non-response rate was still around 30%.

One could argue that the rise in non-response is due to a (drastic) change in the behaviour of potential respondents in response to the Covid-19 shock. It is more likely, however, that it has been caused by a shift in the conduct of Statistics Canada's activities, as indicated by the *timing* of the increase—clearly concomitant to the start of the first national shutdown—and by its striking *persistence*. As of March 2020, federal agencies were instructed to adopt telework for the majority of their employees, i.e. federal workers were required to work from home whenever and wherever possible. Statistics Canada followed suit by suspending its field activities. This represented an important change in the data collection process: interviews could no longer carried out in person, and as such, the LFS had to rely on its other two modes of collection (phone and online) for all interviews.

At first sight, one might think that the changes in collection methods were not dramatic (see figure 2), since the vast majority of interviews were being conducted by phone or online—even before the pandemic (92.8% in 2019). However, a bit more context is required. The LFS has three modes of interview—in person, by phone, and online. In normal times, in-person

[11] There was much commonality in how provinces responded to the outbreak, with more draconian measures being imposed when cases counts became too high (e.g., shutting down in-person schooling). There were, however, some regional differences. Atlantic Canada, for example, was quicker to impose inter-provincial travel restrictions, and Quebec was the only province to impose a nightly curfew.

interviews are conducted for a significant minority of birth interviews (34.5% in 2019).[12] It is also the fall back approach if phone interviewing is unsuccessful (e.g., the telephone number associated to the dwelling proved to be incorrect).[13] These policies reflect the fact that the in-person mode of collection, although more expensive, is known to be effective for making first contact. As such, the halt of in-person interviews is expected to impact non-response rates. Due to the rotating nature of the sample, there is a potential for a significant stock-flow effect: the increased difficulties in making first contact in the incoming rotation can propagate to the ensuing months of the six-month window. A similar argument can be made for subsequent non-response given that in-person contact can help build a relationship with the household. It then follows that an apparently modest shift in interview modes can potentially generate a high and persistent increase in non-response rates, as seen in figure 1.

The suspension of in-person interviewing may only be part of the story, however. In pre-Covid times, telephone interviews were conducted from call centres. The adoption of telework arrangements at Statistics Canada required dispatching surveyors out of these centres, which could have resulted in efficiency losses. Given that it is a long-running survey, the LFS has experienced many expected and unexpected changes/shocks over the years, which can give us guidance on the impact of this most recent "reorganizational" effect.[14] If past experiences are any indication, such changes do impact non-response rates due to the tight production deadline from enumeration to its public release. Having said this, the effects are of second order importance, and most importantly, temporary in nature as compared to what is observed in Covid times. As an illustration, the Quebec non-response rate reached 11.4% during the January 1998 Quebec ice storm, only to come back to 6.2% (a more typical rate at that time) in the following month (Statistics Canada 1998). All in all, these facts point toward the idea that the suspension of in-person interviews is the main explanation behind

[12]Historically the LFS followed a warm telephone approach where the birth interview was in-person, and subsequent interviews were over the phone. Telephone First Contact (TFC) was introduced in 2004 whereas some birth interviews were to be done by phone from a centralized call centre, conditional on a telephone number being available. It should be noted, however, that the LFS had started using telephone birth interviews for those living in high-rises in 1990 and there is mention that it was also used in cases where there was no contact after a personal visit (Sheridan et al. 1997). As such, telephone birth interviews had started growing in importance even before the advent of TFC, i.e. 34.6% of all birth interviews were conducted by phone in 2003. TFC just accelerated its rise. In 2005, after the introduction of TFC, telephone interviewing made up, for the first time, the majority of birth interview (i.e. 55.4%). The number of telephone interviews has continued to creep up after that, reaching 65.5% in 2019.

[13]Other possible reasons for switching to in-person interviewing include: i) two consecutive months of non contact, ii) a telephone number could not be obtained for the dwelling, and iii) the respondent requested in-person interviews (Statistics Canada 2017).

[14]Planned changes have included the introduction of computer-assisted interviewing, TFC, centralized telephone interviewing (at call centres), web-interviewing, and new computer applications, to name but a few. Some unexpected shocks that have occurred include bad weather, an interviewer strike, and computer difficulties.

the persistent surge in non-response rates.

3 Implications of Covid-19 non-response for LFS data

This section analyzes the implications of the surge in non-response since the onset of the outbreak for labour-market estimates that are based on LFS data. The analysis distinguish two types of non-response: *birth* and *subsequent* non-response. Birth non-response refers to an impossibility to establish a first contact with households living in sampled dwellings. Subsequent non-response refers to the inability to reach a household in the current month after contact has been made in a prior month. For instance, consider a household belonging to the incoming rotation of the March 2021 LFS. Suppose first contact is only made in May, i.e. the household could not be reached in either March or April; additionally, assume the household is away in June. In this case, the March and April non-responses are counted as *birth* non-response, whereas the June non-response is counted as *subsequent*.

The reason one should care about the distinction is that the missing information resulting from subsequent non-response can be imputed using the available information of the household. During the household's birth interview, Statistics Canada first collects socio-demographic information of all household members, and then gathers their labour-market detail. In subsequent interviews, only the labour-market information is updated; in the case of subsequent non-response, the socio-demographic information is used, in combination with previous labour-market information to impute the *current* missing information.[15] Obviously, in the case of birth non-response, imputation is not possible, and the household remains out of the LFS sample as long as it cannot be reached by the surveyors.

3.1 The relative importance of birth vs. subsequent non-response

The two types of non-response have different implications for labour-market estimates. Birth non-response could induce additional selection in the LFS sample along observable and non-observable characteristics that is correlated with labour-market outcomes. On the other hand, subsequent non-response resulting in imputation could introduce additional measure-

[15]This is referred to as a "hot-deck" imputation procedure, which works as follows: a group of individual "donors" for whom complete labour-market information has been collected is formed (after excluding outliers with extreme hourly wage values compared to individuals within the same sex/age/province group); then, individuals with missing information are randomly matched with a donor with similar characteristics, and are imputed the information of the donor. In the case of individuals with missing information due to non-response, the matching with a donor is based on a series of socio-demographic variables and information regarding last month's labour-force status and industry. See Statistics Canada (2017) for a more detailed description of the imputation procedure.

ment or classification error. It is therefore important to assess the relative importance of the two non-response causes, and how they have been affected by the Covid-19 disruptions.[16]

We cannot directly attribute a share of non-response due one motive or the other, because we do not have information on the number of individuals living in non-respondent households; recall non-response rates are defined at the household level. However, we know from our restricted-use LFS files the number of individuals with fully imputed information, i.e. the number of individual whole record imputation (WRI) cases. This information allows us, with the help of some mild identifying assumptions, to disentangle the two types of non-response.[17] For our empirical analysis—when relying on the restricted-use LFS files—we restrict our attention to civilian workers aged 20 to 64, and exclude those living in the territories. We impose the age restrictions so as to compare our findings with existing evidence regarding the Covid-19 shutdown (e.g., Lemieux et al. (2020), Jones et al. (2020)).[18] It should be noted that the main conclusions of our paper hold true when the age restrictions are loosened. An upper age limit is needed, however, for the part of the analysis where we identify subsequent non-response using WRI information. This is the case because respondents 70 years and above are only interviewed once, and all their labour market information are automatically carried forward into the remaining months of the six-month window.[19]

The (blue) solid line in figure 3 tracks over time the sample that includes WRI cases—the unrestricted sample observed by researchers. Assuming that in the absence of Covid, the sample would have remained roughly similar beyond February 2020 (i.e. around 70,000 as it has been for all of 2019), the drop beyond this threshold approximates the number of individuals lost due to birth non-response. Following this, one can conclude that around 10,000 individuals are absent from the June 2020 sample due to the Covid-related increase in birth non-response. In December 2020, the number stood at approximately 15,000. The (red) dashed line shows the sample size where WRI cases are excluded; it only accounts for individuals that were interviewed in the current month. Assuming again that this restricted sample would have stayed more or less the same without Covid, the drop in the dashed line reflects Covid-induced birth *and* subsequent non-response. In both June and December,

[16]Moreover, the distinction may affect the choice of public-use versus master files. The variable that identifies individuals with fully imputed labour market information is suppressed in the public-use files. Public-use files users are therefore blind to changing patterns in subsequent non-response. They cannot, for example, carry-out robustness checks to see if their findings are sensitive to the inclusion of individuals whose labour market information was fully imputed.

[17]Indeed, note that birth and subsequent non-response rates are not available to researchers.

[18]Having a lower bound of 20 years of age also means that one can treat individuals who did not complete high school as dropouts—a group for which non-response is typically a problem.

[19]It should be noted that additional restrictions must be imposed when constructing mini-panels for the longitudinal analysis component of our paper. See Online Appendix A.1 for detail.

the decline was about 12,000. As such, the majority (if not all) of non-response cases for these months can be attributed to failed birth interviews. The same holds true for the other months of 2020.

Although of second order importance, the temporary rise in subsequent non-response should nonetheless be of concern to researchers. To illustrate this point, we present in Figure 4 WRI cases as a percentage of the full sample. Before the pandemic, it hovered in the 8-9% band (except for July where it was slightly higher); after March 2021, it rose sharply, reaching 14.1% in June, and only fell (and stayed) below 10% as of October 2020. To put these numbers into perspective, over the March to September 2020 period, the labour market information of one out of eight individuals was fully imputed. For the last two months of our sample (i.e. December 2020 and January 2021), the subsequent non-response was back to historical level with approximately one out of twelve being fully imputed.

Understanding whether birth non-response is mainly an incoming issue, or whether it persists throughout the six-month window, is of particular relevance to LFS panel users.[20] Figure 5 plots the size of the incoming rotation over time. It has certainly dropped due to Covid, especially between February and March, where it decreased by around 3,500 individuals. In the following months, the decline, albeit more tempered, still ranged from 2,000 to 3,000. Although substantive—given that the incoming sample stood at approximately 11,000 in February 2020—these drops can only account for a minor part of the total decline in the sample. This means that LFS surveyors have been, month after month, consistently unable to make first contact with the same households since the advent of the Covid epidemic: a subset of the households have been harder to reach not just in their incoming month, but also in ensuing months, and as such, inducing a gradual increase in non-response rates.

If one imposes additional structure, i.e. assumes that the household non-response rates shown in figure 1 also apply to individuals 20 to 64 year of age, one can quantify more precisely the importance of each type of non-response. Columns (2) and (3) of table 1 show how birth and subsequent non-response have evolved over time, whereas columns (4) and (5), provide a further breakdown of birth non-response by whether it occurs in the first month, i.e. in the incoming rotation, or continues into the latter months of the six-month window.[21] Table 1 clearly shows that prior to the pandemic, birth non-response was of secondary importance as compared to subsequent non-response. For example, in February 2020, birth and subsequent

[20]If non-response observed in Covid times is predominantly an incoming rotation issue, then researchers that rely on two-month mini panels would be advised to use panels that use any two consecutive months, as opposed to the first two months of the six-month window.

[21]In Online Appendix A.2 we provide a detailed description of how each type of non-response was constructed based on the identifying assumptions mentioned above.

non-response were 3.8% and 8.1%, respectively, The striking feature is the rise in birth non-response in the Covid era. The fact birth non-response for the incoming rotation only grew by a few percentage points confirms our prior claim that first contact difficulties persist throughout the six-month window. We carried out robustness checks where we repeated the analysis of table 1 for less restrictive age groups and also using alternative approaches to measuring the number of individuals that Statistics Canada attempts to contact each month. In all cases, the main patterns discussed above nevertheless remain the same. See Online Appendix A.2 for more information.

All in all, this analysis suggests that the major issue that has plagued the LFS has been a surge in birth non-response, with first contact difficulties persisting beyond the first month of the six-month window.

3.2 In-person interviews and individual characteristics

As discussed in section 2, there is a complementarity between the in-person and phone modes of collection, where in-person interviewing is relied upon when there are difficulties with the phone approach. As such, one might expect self-selection into the LFS in Covid times when field operations were stopped. To investigate this possibility, we estimate the following linear probability model:

$$\Pr(inperson_i = 1|X_i) = \alpha_0 + X_i'\alpha, \tag{1}$$

where $inperson_i$ is an indicator variable equal to one if individual i was interviewed in-person (and zero otherwise). X_i is a vector of socio-demographic characteristics comprising of binary variables for gender, age and education groups, and regions of residence.[22]

Table 2 shows regression results for two different sample periods: March 2019 and July 2019. The choice of 2019—the year prior to the pandemic—is to see whether the probability of being interviewed in 'normal' times is correlated with socio-demographic characteristics. Two different months are chosen to account for possible seasonal patterns (as is observed in the non-response rate themselves).[23] Since the analysis of the previous subsection indicates that high non-response rates are mostly accounted for by the falling number of *birth* interviews,

[22]More precisely, the vector includes a female binary variable, two age dummies (20 to 29 and 50 to 64), three highest educational attainment dummies (dropout, college, and bachelor degree and up), and five regions of residence dummies (East, Quebec, Prairies, Alberta and British Columbia). The reference group therefore consists of males that are 30 to 49 years of age with no more than a high school degree who living in Ontario.

[23]July has historically been characterized by low response rates—suggesting difficulties with reaching households could be reflected in the prevailing interview mode.

we report regression results for this sub-population and for the incoming rotation (a further subset), in addition to that of the full sample. Finally, the regressions are unweighted as we are interested in estimating the share of in-person respondents within the sample itself (as opposed to what would be done in the case of estimating population moments).

The findings of table 2 indicate that the interview mode is non-random, especially for *birth* samples. The youngest and low-educated are more likely to be interviewed in person: in March 2019, i.e. just one year before the onset of Covid, individuals 20 to 29 years of age were 5.8 percentage point more likely to do an in-person interview than the 30 to 49 year olds. For dropouts, the likelihood was 4.9 percentage point higher than for individuals whose highest level of educational attainment was a high-school degree. In addition, the 50 to 64 year olds and those with a bachelor degree were less likely to be interviewed in person. The results are overall similar for July 2019; the most notable difference is the greater magnitude associated with the dropout coefficient for this month. There are also important disparities across provinces, with Albertans being the most likely to be interviewed in person (30 additional percentage points compared to Ontarians)—which would suggest distinctly different operating procedures across provinces.

When focussing on the *incoming* rotation, the results are essentially the same; there is again a clear pattern of in-person interview probabilities decreasing monotonically with age and education. This similarity is at least in part due to the important overlap between the two samples (all incoming interviews are by construction birth interviews), but this confirms that a non-randomly selected sub-sample of the households need to be reached in person very early on. Columns (1) and (4) of table 2 assess how the interview modes are distributed across individuals in our full cross-section sample. The results are qualitatively the same (and still statistically significant), but much more subdued. Indeed, most of the in-person interviews are conducted in the incoming rotation, since this is in many cases used as a mean to establish initial contact. As such, the variation in interview modes across individuals is diluted when considering the full cross-section sample. Still, the differences across characteristics are significant even when one considers this sample.

How are these patterns linked with Covid non-response issues? Recall that in-person interviews play a crucial role in reaching households that cannot be contacted by phone. It follows that the fact that this mode of interviewing is more likely to be used for low-educated and younger individuals, suggests that they represent sub-populations that have been especially hard to reach due to shutdown policies accompanying the epidemic—and the resulting suspension of in-person interviews. That interview modes are correlated with characteristics in the birth and the full cross-section samples indicates that non-response might introduce biases in estimates based on incoming samples, but also those based on the full LFS sample. Crucially, the fact that youths and the low-educated seem to have been harder to reach is not insignificant, since it has widely documented that these two groups have been—by far—among the hardest hit in terms of labour-market outcomes (e.g., Lemieux et al. (2020)). Obviously, the remaining question is to what extent can weights mitigate these issues, a question what we will shortly address.

3.3 *Birth* non-response and selection

Selection along socio-demographic characteristics. To further delve into potential selection issues, we explore how Covid-19 has impacted the socio-demographic structure of the LFS. Figure 6 presents unweighted sample shares by gender, age and education groups, and region of residence, for the incoming rotations over the January 2019 - January 2021 period. Analyzing the structure of the incoming group is restrictive, but given the documented difficulties with first-contact interviews, it will provides a clear idea of how the sample of new entrants (to the survey) changed with the advent of Covid-19. The focus is on unweighted proportions as to examine the structure of the sample itself; when discussing the implications for labour-market estimates, we make use of weights.

Figure 6 shows a break in March 2021 along age and education lines. Rotations that started entering as of March 2021 are older and more educated than pre-Covid cohorts. The figure shows a clear drop in the proportion of individuals 20 to 29 years of age (around 20% in February vs. 16% in March), and a spike in the proportion of those aged 50 to 64 (36% vs. 40%). We see as well a decline in the proportions for the high-school and dropout groups, and an increase in the tertiary education groups (college and university). These findings are indicative of a significant and persistent impact of Covid-19 on response behaviours, consistent with the idea that the suspension of in-person activities created difficulties with initial contacts that have generated persistently high non-response rates. Moreover, the break in the trends of the age and education sample shares mirror well the documented facts of the previous section, which shows non-random assignment of interview modes along these two dimensions—with in-person interviews being more likely for the youngest, and the low-education individuals.

In figure 7, we look at the same unweighted shares but for the full sample (i.e. all rotations), to assess how first contact difficulties propagate through the full sample. As one might have expected, one can observe a gradual change in sample structure starting in March 2021; given the rotating panel of the design, it takes a few months to get a critical mass of new entrants. This is especially visible along the age-group dimension: the 20 to 29 age group share has been monotonically decreasing between March and June 2020, whereas the proportion of those 50 to 64 has been following the opposite trajectory. 50 to 64-year-olds account for 40% of the sample in June and onwards, vs. 36% in February (20 vs. 17% for youths). We observe a similar pattern along the education dimension, with a progressive increase in the share of highly-educated individuals in the sample.

We complement the graphical investigation with a regression analysis. It allows us to carry out a series of statistical tests, and most importantly control for confounding factors. We do so by estimating the following linear probability model

$$\Pr(d^m_{i,t} = 1|X_{i,t}, d^{2020}_{i,t}) = \beta^m_0 + X'_{i,t}\beta + \gamma^m_0 d^{2020}_{i,t} + (d^{2020}_{i,t} \times X'_{i,t})\gamma^m, \tag{2}$$

where $d^m_{i,t}$ is an indicator variable equal to one if individual i in year t is observed in calendar month m (see below) instead of February, i.e. the month just before the onset of the epidemic in Canada; X is the vector of socio-demographic variables as defined in equation (1); finally, $d^{2020}_{i,t}$ is an indicator variable taking the value one if the individual is observed in year $t = 2020$.

We estimate equation (2) for different sub-samples, i.e. using data for February and month $m \in$ {March, May, July, September, November} in 2019 and 2020.[24] Hence,

$$\Pr(d^m_{i,2020} = 1|X_{i,2020}) - \Pr(d^m_{i,2020} = 1|X_{i,2019}) = \gamma^m_0 + X'_{i,2020}\gamma^m \tag{3}$$

captures the effect of the Covid disruption on the probability of appearing in a Covid month instead of February 2020, conditional on characteristics.[25] More specifically, γ^m—the parameter vector of interest—measures how Covid-induced selection into non-response varies across observable characteristics.[26]

Table 3 shows the unweighted results for the incoming rotation. Consistent with the graphical analysis (i.e. the unconditional findings) one still observes age and education profiles that

[24] Said differently, we first focus on February and March data for 2019 and 2020, and estimate the conditional probability of appearing in March. We then use February and May data (again for 2019 and 2020) to examine the conditional probability of being observed in May. The regressions where m represents the months of July, September, and November are similarly defined. It should be noted that by using 2019 data we can account for seasonal patterns in non-response.

[25] Unless a substantial change in the data collecting process was introduced between 2019 and 2020. To the best of our knowledge, no such change have occurred.

[26] If the Covid effect on the probability of being observed was truly random, then γ^m=0 and $\gamma^m_0 < 0$ (for $m \in$ {March, May, July, September, November}), with the negative intercept reflecting the increase in non-response observed in the Covid-era.

persist over time—even when controlling for confounding factors. These effects tend to be economically and statistically significant. For instance, the probability 20 to 29 year olds appear in the March incoming rotation (instead of February) is 4.7 percentage points (p.p.) lower in 2020 than in 2019, relative to the 30 to 49 year old group. Even more than six months after the initial shutdown, the youth age group has a 3.6 p.p. lower probability of appearing in the September incoming rotation. If one focusses on the education dimension, one sees that in Covid times, individuals with a bachelor (or more) are 3.5 p.p. more likely to be a March entrant, and 3 p.p. more likely to be a September entrant (relative to the high-school individuals).

To see how this incoming-rotation 'entry' selection pattern affects the structure of the full LFS sample, we report in table 4 unweighted regressions results for all rotations. Again, the estimates suggest a significant change in the structure of the sample along age and education dimensions. Consistent with the idea that non-response is primarily a birth issue, the changes in the sample structure come across gradually: there is no discernible effect for March, but when focussing on later months, the full LFS sample is clearly moving away from youths and low-educated individuals. As of November—*eight* months after the start of the epidemic—individuals 20 to 29 years of age are 2.8 p.p. less likely to be present in the sample than in the previous year (in comparison to the prime-age reference group); even more striking is the 4 p.p. lower probability of being part of the LFS sample for dropouts.

Finally, we complement the analysis by observing the impact of Covid on the structure of *balanced panels*. In the online appendix, we show the results for 'new-entrant' mini panels (i.e. individuals are observed in the first two month of the six-month window) and the 'full-rotation' panels (i.e. individuals are observed in any two consecutive months). To make our panel analysis comparable with the cross sectional one, the referenced month (i.e. February or month m) is taken to be the first period of the mini-panel. For example, the panel equivalent of being observed in March (as opposed to February), is the probability of being observed in the March-April panel (as opposed to the February-March panel).

Unsurprisingly, the results for the new-entrant panel are similar to those of the incoming-rotation, and this for two related reasons: first, the new entrant mini-panel is, by construction, very similar to the incoming rotation; in both cases, the person must be part of the incoming rotation as of the referenced month. Second, the added panel restriction that the individual be also observed in the following month is not overly restrictive in Covid times since we know that subsequent non-response is not the driving force behind the important rise in non-response. On the other hand, the results for 'full-rotation' panels are more in line with those of the full cross-sectional sample. This holds true because in both cases the majority of individuals are not new to the sample. Having said this, one does lose a full rotation with the panel approach as those that are in the outgoing rotation group as of the first month of the panel cannot reappear in the second month.

Hence, due to the nature of the non-response and selection process—primarily driven by the new LFS entrants—researchers might expect to face different types of selection patterns depending on which panel design they decide to adopt. Essentially, researchers using new-entrant panels should expect selection as of the beginning of the pandemic while those opting for full-rotation panels should instead be worried about selection being an issue later in the year, i.e. in the summer and fall months. Moreover, the full-rotation panel features clear and persistent selection changes for the low educated and across the full age distribution, whereas, in the new-entrant panel, the Covid effect is most apparent for the youths. Regarding the latter aspect, the fact that the full-rotation panel selection pattern looks similar to that in the cross-sectional sample might be good news: since the LFS weights construction is based on cross-sectional information, we might expect these weights to mitigate selection issues in the full-rotation panels. We come back to this question shortly.

Selection along employment and job characteristics. We now analyze how employment and job characteristics correlate with non-response behaviours. A substantial literature has documented important heterogeneity in the impact of Covid on labour-market outcomes across industries and occupations in Canada and elsewhere (e.g., Lemieux et al. (2020); Dingel & Neiman (2020); Baylis et al. (2020)). It is therefore key to investigate whether workers who have been more exposed to the shock due to the characteristics of their job are more or less likely to be selected in the LFS sample in March 2020 onward.

Note that here, we cannot repeat the exercise of analyzing the change in the sample structure using 2019 and 2020 data (i.e. pre-Covid and Covid-era data). Indeed, job characteristics have been directly impacted by Covid beyond any change due to non-response; the Covid-led recession, for example, had a direct impact on the proportion individuals that are still employed and the type of job they hold. One possibility would be to look for evidence of selection based on the labour-market situation just *before* the epidemic—of those observed during the pandemic. However, since we are mainly interested in understanding *birth* non-response behaviours, this would require having retrospective information for new entrants in the LFS. Since such information is rather limited in the LFS, we cannot easily follow this approach.

We examine instead whether job characteristics that have been documented to be correlated with the intensity of Covid's impact on individual outcomes are correlated with the mode of interview before Covid (i.e. in 2019). In particular, we explore whether in-person respondents differ systematically from remotely-interviewed individuals, conditional on their job and socio-demographic characteristics. To do so, we estimate an augmented version of the linear probability model shown in equation (1), a version that includes job characteristics information as additional explanatory variables. The added variables consist of dummies for the employment status, quartile of the individual weekly salary, length of job tenure, and telework potential of the individual's occupation.[27] Considering individuals' salary is motivated by studies that have documented a higher adverse impact of Covid on low-earning workers (e.g., Lemieux et al. (2020)) and accounting for job tenure is prompted by our previous results in Brochu et al. (2020) that suggest higher employment losses among low-tenure workers (controlling for age). Finally, the introduction of a telework variable follows the studies that have documented a more negative impact on individuals in occupation with low telework potential (e.g., for the U.S.: Mongey et al. (2021); for Canada: Baylis et al. (2020)).[28] It should be noted that we are interested in examining the presence of heterogeneity after controlling for characteristics that are taken into account in the design of the LFS weights (shortly discussed). As such, we also introduce in the regression a set of dummies for the one-digit industry group of employed workers.[29]

The results are shown in table 5. There are striking differences along labour-market characteristics. In March 2019, low-salaried and low job-tenured workers, and those working in occupations with low potential for telework are more likely to be interviewed in person. Once again, the magnitude of these differences is higher for birth and incoming-rotation interviews than for the full sample (but the significance tends to be lower, probably due to due smaller sample size). For instance, in birth and incoming interviews, individuals in the second salary quartile are between 3 and 4 p.p. more likely to be interviewed in persons than those in the third quartile. Strikingly, in person-interviews are between 6 and 7 p.p. more likely for the employed with less than one year of tenure, than for the workers in more stable,

[27] For job tenure we include two binary variables: one for jobs with less than one year of tenure with the same employer, and another for jobs with one to less than three years of tenure. It should be noted that salary information is only collected of employees in the LFS.

[28] We rely on the telework feasibility indicator constructed by Deng et al. (2020), which is based on the one developed by Dingel & Neiman (2020). It is adapted to the Canadian standard NOC 2011 occupation classification system (from the U.S. SOC classification system) using the Social Analysis and Modelling Division of Statistics Canada crosswalk. See Dingel & Neiman (2020) and Deng et al. (2020) for additional detail.

[29] It should be noted that the LFS also gathers some job characteristics of the last job for those that are out of work, but were employed in the last year. We are interested in job characteristics of those that are employed as of the current month. As such, job characteristics only enter the equation as interaction terms—interacted with employment.

long-lasting jobs (i.e. three years of tenure and more). Similar patterns are observed in July. Keep in mind that these regressions control for gender, age, education, provinces, and industries.[30] As such, these differences reflect heterogeneity in response behaviour conditional on the above-mentioned characteristics. Moreover, remember as well that the regressions include industry controls.

Although we cannot directly test for changes in the structure of the LFS sample along pre-Covid job characteristics, we interpret these results as indirect evidence that such changes have been occurring. The systematic differences in the profile of in-person respondents before Covid, combined with the surge in non-response due to the suspension of this type of interviews points to this direction. Furthermore, since there is heterogeneity within key socio-demographic characteristics (gender, age, region) and within industry groups, we should expect that the selection resulting from Covid has not been reflected in the adjustment of LFS weights. The next section includes a discussion of this particular issue.

4 Implications for estimating the impact of Covid-19

4.1 Non-response and LFS weights

This section examines the extent to which the LFS weights mitigate the potential selection biases linked to the Covid-induced surge in non-response. Remember that the LFS imputation procedure relies on information that has been collected during prior interviews. As such, it can only be carried out for cases of subsequent non-response. Households for which no first contact could be made are therefore dropped from the sample. The LFS weights are, however, designed to account for this loss of information.[31]

The LFS weighting methodology has a rich procedure. Key to our study is the "calibration" step, which is performed to ensure that (i) the weighted sample is representative of the entire population along key select dimensions including gender, age, and region of residence—but *not* education; and (ii) it has a consistent structure across select labour-market dimensions over time (i.e. across ensuing months). For the latter, the weights are calibrated to ensure that individuals in the time t sample who also had been observed in $t-1$, have similar $t-1$ labour-market characteristics as those of the full $t-1$ sample.[32] The labour market

[30] In particular, the job-tenure pattern holds when controlling for age at a detailed level (i.e. with a set of one-year age dummies).

[31] They will also account for non-random aspects of the LFS sampling design, i.e. that it follows a stratified sample design.

[32] An example will best illustrate this point. Assume the LFS is calibrating the weights of the August 2020 sample. The LFS focusses on a sub-group of this sample—individuals that also had responded in the

dimensions that are accounted for include: (i) employment, unemployment, and participation (in aggregate and by sex/age groups); (ii) employment by industry (corresponding roughly one-digit NAICS groups); (iii) employment in the public and the private sector.[33]

Since the weight calibration is performed to ensure that the sample is representative of the Canadian population across gender, age, and region, one should not, when weights are applied, be able to detect any changes as of March 2020 along these dimensions. Figures 8 and 9 display weighted sample shares by gender, age, region and *education*, for the incoming rotation and the full sample, respectively. As expected, there is no apparent break in trends along the gender, age and region dimensions (in direct contrast to what was observed for the unweighted shares shown figures 6 and 7). This result stands when controlling for confounding factors, i.e. when we run weighted regressions of equation (2) for the full cross-sectional sample (table 6); the gender, age, and region effects that were clearly apparent in the unweighted regressions essentially vanish with the application of weights. Therefore, the weight adjustments have been mitigating the Covid-induced selection issues for some key socio-demographic characteristics.

It is worth noting that (along these same three dimensions) the weights also seem to considerably mitigate the selection issues for the *panels*.[34] This might come as a bit of surprise since the weights are adjusted and calibrated based on estimates for the entire LFS sample. Indeed, constructing panels puts additional restrictions as it requires information on the same individuals for subsequent months. But, as explained before, since the non-response issue is mostly due to failed birth interviews (that persist through the six-month window) rather than subsequent non-response, the structure of the full-rotation panel samples is similar to that of the full cross-sectional sample (and the new entrant panel to that of the incoming rotation). We see this as good news for researchers that rely on the longitudinal dimension of the LFS.

However—by construction—the weights do *not* correct for other key dimensions, such as education, for which we have documented significant selection. In table 6, the dropout coefficients are still highly significant and have magnitude close to these found for the unweighted

prior month (i.e. July 2020)—and compares their labour market outcomes as of July 2020 to those of the full July 2020 sample. The August 2020 weights are then adjusted/calibrated to account for any systematic differences.

[33]Specifically, this calibration procedure consists in minimizing the distance between the "sub-weights" (i.e. the weights obtained prior to the calibration) and the "calibrated" weights, under the constraint that the sample produces consistent statistics across consecutive months, and population-share estimates for key demographic characteristics that are consistent with those from the Census. See Statistics Canada (2017) for a complete description of the weighting procedure.

[34]The tables of results can be found in the Online Appendix.

regressions of table 4. We should expect, therefore, that the weights have not been able to correct for the heterogeneity in job characteristics. Indeed, the previous section presents evidence suggesting heterogeneity in response behaviour conditional on socio-demographic and job characteristics. It now remains to evaluate the implications of the Covid non-response on labour-market estimates, which is the object of the next subsection.

4.2 The Covid-19 labour-market impact across LFS rotations

This section exploits the rotating-panel design of the LFS to examine the implications of selection for estimating the impact of Covid on labour-market outcomes. Indeed, this rotating-panel design, combined with the fact that non-response is mainly associated with birth non-response (and, as such, with incoming rotations), provides variation in selection behaviours that is uncorrelated with labour-market outcomes.

Essentially, our approach consists of comparing the impact of Covid on the individuals assigned to the rotations that entered the LFS *before* the onset of the epidemic in Canada (i.e., before March 2020), with those assigned to the rotations entering the LFS *afterwards*. Indeed, as we have argued so far, being assigned to a rotation that joined the LFS after the onset of Covid (which we call a post-Covid rotation) is associated with a lower likelihood of participating in the survey, presumably due to higher difficulties with reaching new entrants given the suspension of in-person interviews. On the other hand, we should expect the rotation assignment of individuals (for a given month) to be independent of their labour-market outcomes. As such, the differences between the two groups after the beginning of Covid should reflect different selection patterns.

Hence, the key idea behind our proposed approach is that the individuals assigned to pre-Covid rotations have been exposed to a different *selection rule* than those assigned to post-Covid rotations. Before Covid, in-person interviews were available; however, this mode of interview was abruptly abandoned as of March 2020. As such, the post-Covid group have a higher non-response rate due to an exogenous factor—the assignment to a given LFS rotation.

Specifically, we propose the following model

$$E(y_{i,t}|d^r_{i,t}, d^{2020}_{i,t}) = \delta_0 + \delta_1 d^r_{i,t} + \delta_2 d^{2020}_{i,t} + \delta_3 (d^r_{i,t} \times d^{2020}_{i,t}), \qquad (4)$$

where $y_{i,t}$ is an individual labour-market outcome and $d^r_{i,t}$ is an indicator taking the value one if the individual is assigned to a rotation that entered the LFS in March or a later calendar month for year t. Finally, $d_{i,t}^{2020}$ is again defined as an indicator variable that equals one if the individual is observed in the year 2020. We estimate equation (4) for different sub-samples. We first estimate the model using March data for 2019 and 2020. We then repeat the exercise for each of the four subsequent months, separately. It should be noted that we only go up to July because this exercise requires having individuals that are part of rotations that entered the LFS prior to the pandemic. By July, only one rotation still meets this criteria, i.e. the rotation which entered in February.

In particular,

$$E(y_{i,2020}|d_{i,2020}^{r}=1)-E(y_{i,2019}|d_{i,2019}^{r}=1)- \\ \big[E(y_{i,2020}|d_{i,2020}^{r}=0)-E(y_{i,2019}|d_{i,2019}^{r}=0)\big]=\delta_3 \tag{5}$$

estimates the additional effect of Covid that is associated with being assigned to a post-Covid rotation instead of a pre-Covid one. As argued before, the differences in labour-market outcomes across these two groups should reflect heterogeneity coming from their different selection rules. Hence, we interpret the estimated coefficient of δ_3 as reflecting differences due to selection induced by the Covid disruption. Notice that the difference-in-difference design of our model allows us to control for seasonal effects that might generate differences across the rotations independently of Covid (especially during the summer months, where non-response is usually high). In what follows, we consider the ensuing binary dependent variables: employment, employment-at-work, and non-participation. Considering employment-at-work is motivated by the dramatic decline in hours caused by Covid-19 in Canada (e.g., Lemieux et al. (2020)), and by the historical increase in absences from work (e.g., Jones et al. (2020)).

Before proceeding, we discuss the limitations of this approach. The δ_2 coefficient measures the Covid impact on individuals that have been assigned to a LFS sample prior to the pandemic. Yet, the group constituted by these individuals is clearly not exempted from selection, including selection caused by birth non-response. To illustrate this point, consider for example, two individuals (A and B) assigned to the rotation that entered the LFS in February 2020. Say individual A can be reached in February but not in March, whereas B cannot be reached in either month. Then as of March, individuals A and B are subsequent and birth non-respondents, respectively. Since δ_2 essentially estimates the impact on a group where individuals like B are still present, it is still subject to a non-response bias (caused by birth non-response) that might have been reinforced by Covid. Therefore, our approach is not aimed at completely eliminating a potential non-response bias.

However, provided that (i) the bias is higher for individuals selected according to the post-

Covid rule and that (ii) it has the same sign across rotations, the δ_3 coefficient indicates the direction of the bias.[35] Assumption (i) is justified by the results of subsection 3.1, indicating a clear increase in birth non-response driven by the rotations entering the LFS after the start of Covid. The reader might be more skeptical about assumption (ii), but the results of subsection 3.3 suggest that the selection patterns have essentially remained unchanged along key observable characteristics, across all rotations. At the very least, the estimated value of δ_3 should be seen as a test of the null hypothesis that selection due to the suspension of in-person interviews is uncorrelated with the impact of Covid.

Table 7 shows OLS results for equation (4). Columns (1) to (5) report the findings where we impose $d^r = 0$, i.e. it estimates the unconditional impact of Covid across all rotations, for March to July, whereas columns (6) to (10) report the results for the full regression, i.e., it allows the impact to vary across individuals assigned to pre- and post-Covid rotations. In the five months considered in panel A of table 7, the point estimate of δ_3 is always positive. This indicates that the surveyed individuals who have been assigned to a post-Covid rotation experienced *lower* (net) employment losses than individuals who have been assigned to prior rotations. For the two first months following the onset of Covid, the impact is strong—3.5 p.p. for March and 2.6 p.p. for April, as compared to the pre-Covid rotations—and significant at the 1% level. The magnitude and significance gradually diminish over the year, which is not surprising given that our treatment group is by construction less and less contaminated by birth non-response.[36]

The strong and gradually vanishing impact shown by table 7 suggests the presence of an attenuation bias due to birth non-response when estimating the impact of Covid on employment. As such, using the entire cross-section sample might understate the negative impact of the shock. This can be seen by comparing the coefficients of the restricted and unrestricted models (columns (1) to (5) and (6) to (10), respectively). For all the months considered, the estimated impact on the cross-sectional sample is less negative than for the group of pre-Covid LFS entrants only. This is in line with the analysis of our previous sections arguing that the individuals with the highest employment Covid risk (in terms of observable characteristics) have been hardest to contact. The magnitude of these divergences is fortunately moderate. In March, the unconditional cross-sectional employment impact is -3.1 p.p., vs.

[35] Implicitly, this requires that the bias due to subsequent non-response and imputation does not determine the sign of the bias conditional on the rotation assignment. This is consistent with the fact that subsequent non-response is a minor issue as compared to birth non-response in Covid times.

[36] Recall that δ_3 measures the impact on the group of individuals assigned to post-Covid rotations. In this group, the share of individuals assigned to an incoming rotation is decreasing over time: in March, all the observed post-Covid rotations belong to an incoming rotation; in April, one half (approximately) of the individuals is in an incoming rotation, whereas the other half is in an ongoing rotation; etc.

-3.7 p.p. for the restricted sample; for the other months, the difference is at most equal to 1.1 p.p.

We find similar results for the impact of Covid on the employment-at-work and non-participation probabilities: these probabilities have been less negatively affected for individuals in the March-onward rotations (panels B and C of table 7). The pattern is similar to that of the employment estimates: the δ_3 coefficient has a high magnitude and significance in March and April (less so for the employment-at-work probability than for participation, however), but the impact gradually vanishes. Once again, the point estimates for the impact on the cross-section suggest a lower impact than for individuals assigned to a pre-Covid rotation, for all months. But, as for the employment estimates (table 7), these differences remain moderate.

We now run placebo regressions using 2018-2019 data, for both employment and non-participation.[37] For these years, the employment and participation probabilities tend to be *lower* for the LFS entrants of March-2019 and after. In March and April 2019, however, there is no significant rotation effect. But for the late spring-early summer months, the effect is substantial and significant. Here is a possible interpretation. Between 2018 and 2019, there have been aggregate employment gains. These gains have been presumably heterogeneous in the population, and one might expect that the most mobile groups have had the higher employment gains. Since mobility is relatively high in May-July, the most economically active individuals might have been harder to reach, especially those in incoming rotations. In sum, we might expect that when the economy is growing, birth non-response induces selection out of the LFS for the high-employment probability individuals, especially around the summer. The same reasoning can be made for participation, with the high-participation individuals being less represented in the incoming rotations of the spring and summer.

Consistent with this mobility-based interpretation are our results for OLS estimations for specific age groups (20 to 29 and 50 to 64 year olds). Conditional on belonging to these two groups, there is no rotation effect for 2018-2019. In other words, the effect seems to disappear when we condition on age. This suggests that the rotation difference for 2018-2019 can be explained by an age-related pattern, with youths becoming harder to reach in the spring-summer due to their higher mobility. Moreover, we find that for 2019-2020, the post-Covid rotation effect is very weak for this groups. An explanation is that there has been low heterogeneity in non-response behaviour within the youth group. The detailed results are shown in the supplementary appendix.

[37] The placebo findings are available in the online appendix.

However, for older workers, the picture depicting the Covid labour-market impact is strikingly different as compared to youths. Table 8 shows the results for this age group, for 2019-2020. In March and April 2020, the employment effect of Covid on 50-64–year-old workers assigned to incoming rotations of March and afterwards is much lower than for the pre-Covid LFS entrants. In March, the employment probability is 4.9 p.p. higher for the LFS entrants than for the incumbent, and 4 p.p. higher in April. The picture looks similar for participation. LFS entrants have been 3.1 p.p. less likely to not be in the labour force in March, and 3.3 p.p. in April. For this group, the point estimates of the coefficient of interest are much higher than for the entire population (table 7). This suggests the presence of a substantial bias due to selection for the 50-64 year-old individuals.

5 Discussion

This paper shows that the increase in non-response due to the propagation of Covid-19 in Canada, coupled with social distancing policies, has not been random. It appears that subpopulations that were vulnerable to the Covid economic shock have been harder to reach since the onset of the pandemic. Most of the increase in non-response, and the associated selection of individuals into LFS samples, can be plausibly attributed to the interruption of in-person interviews.

Since those most exposed to the Covid economic shock have also been the hardest to reach, we might expect estimates of the labour-market impact of Covid to underestimate the true decline in employment and participation. Consistent with this premise is our analysis showing that individuals assigned to a pre-Covid rotation—the ones that entered the LFS prior to March 2020—were less affected than those assigned to a post-Covid rotation. Provided that the assignment of individuals to rotations (when conditioning on a survey month) is uncorrelated with labour-market outcomes, we interpret such differential impacts as reflecting selection. It should be noted that by construction, the LFS weights do not mitigate selection along some key dimensions such as education, the length of job tenure, and the telework potential of occupations.

All-in-all, the results suggest that existing studies might have underestimated the negative labour-market impact of Covid-19. However, the rise in non-response, although large by LFS standards, must be taken into perspective. Given that pre-Covid non-response rates were in the low teens, a resultant non-response rate of 30% more than rivals the rates observed in most surveys—even by pre-Covid standards. As such, we believe that the LFS should still be seen as being a reliable data source for analyzing the economy in Covid times, especially given its relatively high frequency and timely availability to researchers. Furthermore, there are no indications that the non-response issue has dramatic implications for the Covid-19 labour-market narrative that has been proposed, at least when it comes to looking at aggregate outcomes. The differences that we find between individuals assigned to pre- and post-Covid rotations are statistically significant but of moderate magnitude. It might be that there is a substantial bias associated with the estimates when using the full sample (where both groups are present), or it might be that the true bias is modest in size. We lean towards the latter interpretation, as in-person interviews represented a low share of total interviews before the pandemic.

Another reason to be optimistic is that the structure of the panels in terms of socio-demographic characteristics has remained similar to that of cross-section samples during Covid times. Researchers conducting longitudinal analysis should, therefore, feel reassured that panels have not been contaminated by selection issues that were more severe than for the cross-section samples. Essentially, this is because of the relatively low importance of subsequent non-response, which, if high, would have generated excess panel attrition. The low subsequent non-response rates should also alleviate concerns with using public-use files instead of the restrictive-access master files. Indeed, the public-use files have no data imputation information, as opposed to the master files. At first glance, one should expect that access to this information is required for producing robust analysis using Covid-time data. However, the fact that birth rather than subsequent non-response is the main issue tends to mitigate the relative benefits of master files in this regard. Our previous work showing minor differences in key labour-market statistics across LFS samples with and without individuals with fully imputed information (Brochu et al. (2020)) points in the same direction.

The following caveat must be highlighted. Estimates based on specific sub-samples might need to be interpreted with care since the selection issue is presumably non-uniform across individuals—particularly for the months immediately following the first (i.e. spring 2020) lockdown. For instance, the incoming rotations are more affected by the Covid-induced selection issue. Hence, it might be preferable to rely on full-rotation panels than on the new-entrant panels for conducting longitudinal analysis. Moreover, our regressions suggest substantial differences in the Covid impact between rotations, when focussing on the oldest workers. But in the end, when these results are put in perspective with the tremendous disruption caused by Covid in Canada, one can argue that these issues are moderate.

It would be informative to reassess the economic impact of Covid when the relevant administrative datasets will be available to researchers. Through a comparison of survey and administrative data estimates, one could more precisely evaluate the extent to which a pandemic-driven disruption such as the one currently experienced—and which might eventually reoccur—affects the collection of survey data. Moreover, this could shed additional light on the heterogeneity of the Covid impact. As suggested by our analysis, the magnitude of the non-response bias might vary across subgroups. As such, the true distributional impact of Covid might not be fully revealed by survey datasets. A remaining question is to what extent the imputation procedure has been mitigating the subsequent non-response issue. Even if subsequent non-response rates have been modest compared to birth non-response, it might be important to assess its impact given the unprecedented surge in the prevalence of 'unusual' labour force status, such as the absences from work or temporary unemployment. Again, having access to administrative data will eventually shed light on this question.

Our results further suggest that future work could be done to improve weight adjustment to deal with non-response issues. More generally, it might be important to improve our strategies for dealing with the secular increase in non-response rates observed in Canada and elsewhere, especially since the nature of jobs might be evolving quickly—due to technological changes and the expected increased prevalence of telework compared to pre-Covid times—and introduce additional heterogeneity in labour market outcomes. This might prove useful for facing future disruptive events having highly heterogeneous labour-market impacts as was the case for the Covid shock.

References

Achou, B., Boisclair, D., D'Astous, P., Fonseca, R., Glenzer, F. and Michaud, P.-C. (2020), 'Early impact of the COVID-19 pandemic on household finances in Quebec', Canadian Public Policy **46**(S3), S217–S235.

Barrett, G., Levell, P. and Milligan, K. (2014), A comparison of micro and macro expenditure measures across countries using differing survey methods, in C. D. Carroll, T. F. Crossley and J. Sabelhaus, eds, 'Improving the Measurement of Consumer Expenditures', NBER Chapters, National Bureau of Economic Research, Inc, pp. 263–286.

Baylis, P., Beauregard, P.-L., Connolly, M., Fortin, N., Green, D. A., Gutierrez Cubillos, P., Gyetvay, S., Haeck, C., Molnar, T. L., Simard-Duplain, G., Siu, H. E., teNyenhuis, M. and Warman, C. (2020), The distribution of COVID-19 related risks. NBER Working Paper, No. 27881.

Béland, L.-P., Brodeur, A., Mikola, D. and Wright, T. (2020), The short-term economic consequences of COVID-19: Occupation tasks and mental health in Canada. IZA Discussion Paper Series, No. 13254.

Béland, L.-P., Fakorede, O. and Mikola, D. (2020b), 'Short-term effects of COVID-19 on self-employed workers in Canada', Canadian Public Policy **46**(S1), S66–S81.

Bernhardt, R., Munro, D. and Wolcott, E. (2021), How does the dramatic rise of CPS non-response impact labor market indicators? GLO Discussion Paper, No. 781.

Blit, J. (2020), 'Automation and reallocation: Will COVID-19 usher in the future of work?', Canadian Public Policy **46**(S2), S192–S202.

Brochu, P. (2021), 'A researcher's guide to the Labour Force Survey: Its evolution, and the choice of public-use versus master files', Canadian Public Policy **47**(3).

Brochu, P., Créchet, J. and Deng, Z. (2020), Labour market flows and worker trajectories in Canada during COVID-19. CLEF Working Papers Series, No. 32.

Brochu, P. and Green, D. A. (2013), 'The impact of minimum wages on labour market transitions', Economic Journal **123**(573), 1203–1235.

Bryan, K. A., Buajitti, E., Rosella, L. C. and Goel, V. (2021), 'The economic and long-term health consequences of Canadian COVID-19 lockdowns', Canadian Public Policy **47**(2), 281–300.

Deng, Z., Morissette, R. and Messacar, D. (2020), Running the economy remotely: Potential for working from home during and after COVID-19, StatCan COVID-19: Data to Insights for a Better Canada, no. 26. Statistics Canada Cat. no. 45280001, Ottawa.

Dingel, J. I. and Neiman, B. (2020), 'How many jobs can be done at home?', Journal of Public Economics **189**(C).

Forsythe, E., Kahn, L. B., Lange, F. and Wiczer, D. (2020), 'Labor demand in the time of COVID-19: Evidence from vacancy postings and UI claims', Journal of Public Economics **189**(C).

Gallacher, G. and Hossain, I. (2020), 'Remote work and employment dynamics under COVID-19: Evidence from Canada', Canadian Public Policy **46**(S1), S44–S54.

Jones, S. R. G., Lange, F., Riddell, W. C. and Warman, C. (2020), 'Waiting for recovery: The Canadian labour market in June 2020', Canadian Public Policy **46**(S2), S102–S118.

Jones, S. R. G., Lange, F., Riddell, W. C. and Warman, C. (2021), Canadian labour market dynamics during COVID-19. CLEF Working Papers Series, No. 37.

Jones, S. R. G. and Riddell, W. C. (2019), ‘Unemployment, marginal attachment, and labor force participation in Canada and the United States’, Journal of Labor Economics **37**(S2), 399–441.

Koebel, K. and Pohler, D. (2020), Labor markets in crisis: The causal impact of Canada’s COVID-19 economic shutdown on hours worked for workers across the earnings distribution. CLEF Working Papers Series, No. 25.

Kroft, K., Lange, F., Notowidigdo, M. J. and Tudball, M. (2019), ‘Long time out: Unemployment and joblessness in Canada and the United States’, Journal of Labor Economics **37**(S2), 355–397.

Lemieux, T., Milligan, K., Schirle, T. and Skuterud, M. (2020), ‘Initial impacts of the COVID-19 pandemic on the Canadian labour market’, Canadian Public Policy **46**(S1), S55–S65.

Lluis, S. and McCall, B. (2019), ‘Employment and job search implications of the extended weeks and working while on claim pilot initiatives’, Canadian Public Policy **45**(2), 129–172.

Messacar, D., Handler, T. and Frenette, M. (2021), ‘Predicted earnings losses from graduating during COVID-19’, Canadian Public Policy **47**(2), 301–315.

Meyer, B. D., Mok, W. K. C. and Sullivan, J. X. (2015), ‘Household surveys in crisis’, Journal of Economic Perspectives **29**(4), 199–226.

Milligan, K. and Schirle, T. (2019), ‘Push and pull: Disability insurance, regional labor markets, and benefit generosity in Canada and the United States’, Journal of Labor Economics **37**(S2), 289–323.

Mongey, S., Pilossoph, L. and Weinberg, A. (2021), Which workers bear the burden of social distancing? NBER Working Paper, No. 27085.

Montenovo, L., Jiang, X., Lozano Rojas, F., Schmutte, I. M., Simon, K. I., Weinberg, B. A. and Wing, C. (2020), Determinants of disparities in COVID-19 job losses. NBER Working Paper, No. 27132.

Qian, Y. and Fuller, S. (2020), 'COVID-19 and the gender employment among parents of young children', Canadian Public Policy **46**(S2), S89–S101.

Rivera Drew, J. A., Flood, S. and Warren, J. R. (2014), 'Making full use of the longitudinal design of the Current Population Survey: Methods for linking records across 16 months', Journal of Economic and Social Measurement **39**(3), 121–114.

Rothbaum, J. and Bee, A. (2021), Coronavirus infects surveys, too: Nonresponse bias and the coronavirus pandemic. SEHSD Working Paper Number 2020-10.

Schirle, T. (2015), 'The effect of universal child benefits on labour supply', Canadian Journal of Economics **48**(2), 437–463.

Sheridan, M., Drew, D. and Allard, B. (1997), Response rate and the Canadian Labour Force Survey: Luck or good planning?, in 'Proceedings of Statistics Canada Symposium 96, Nonresponse Errors', pp. 67–75.

Statistics Canada (1998), Labour Force Survey: Quality Report Surveys, 0198 to 0698, Statistics Canada Cat. no. 71-0004, 1998, Ottawa.

Statistics Canada (2017), Methodology of the Canadian Labour Force Survey, Statistics Canada, Cat. no. 71-526-X, December 2017, Ottawa.

U.S. Bureau of Labor Statistics (2020), The impact of the coronavirus (COVID-19) pandemic on the employment situation for July 2020. Available at https://www.bls.gov/covid19/employment-situation-covid19-faq-july-2020.htm.

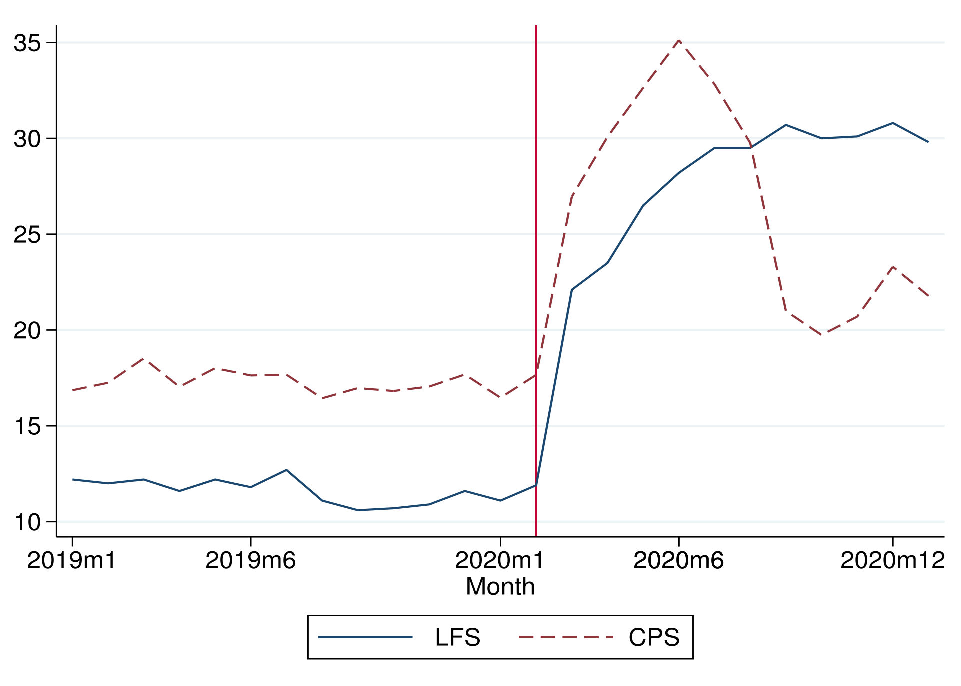

Figure 1: Monthly household non-response rates (%). Source: LFS - Statistics Canada; IPUMS-CPS (`https://cps.ipums.org/cps/covid19.shtml`, accessed June 10, 2021). Share of households in the monthly LFS with no available information. The red vertical line indicates February 2020. The LFS data includes all LFS households in the ten Canadian provinces.

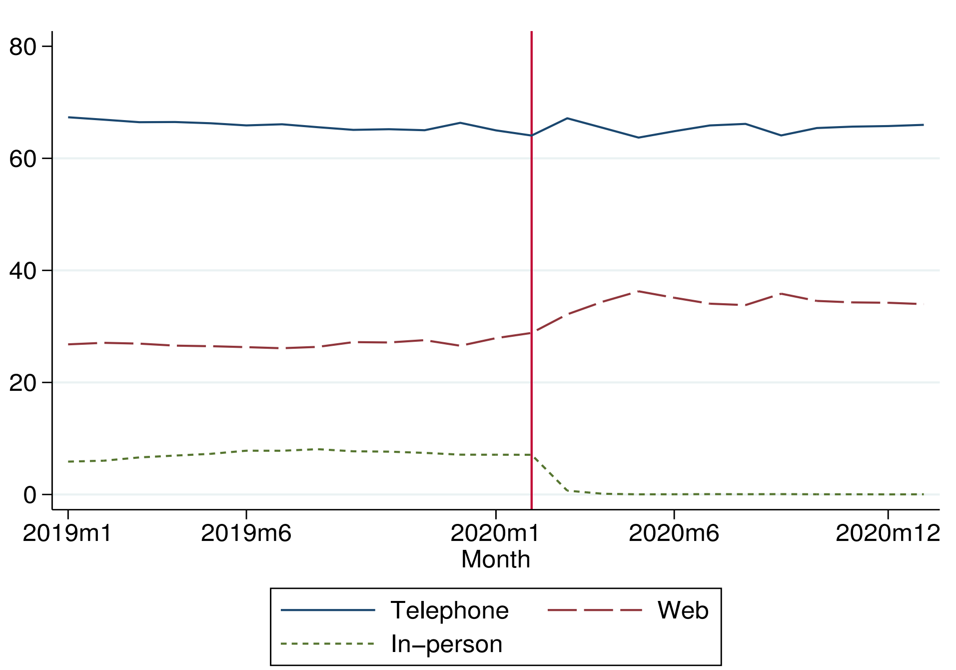

Figure 2: Proportions of interviews by mode of collection (%). Source: LFS and authors' calculations. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

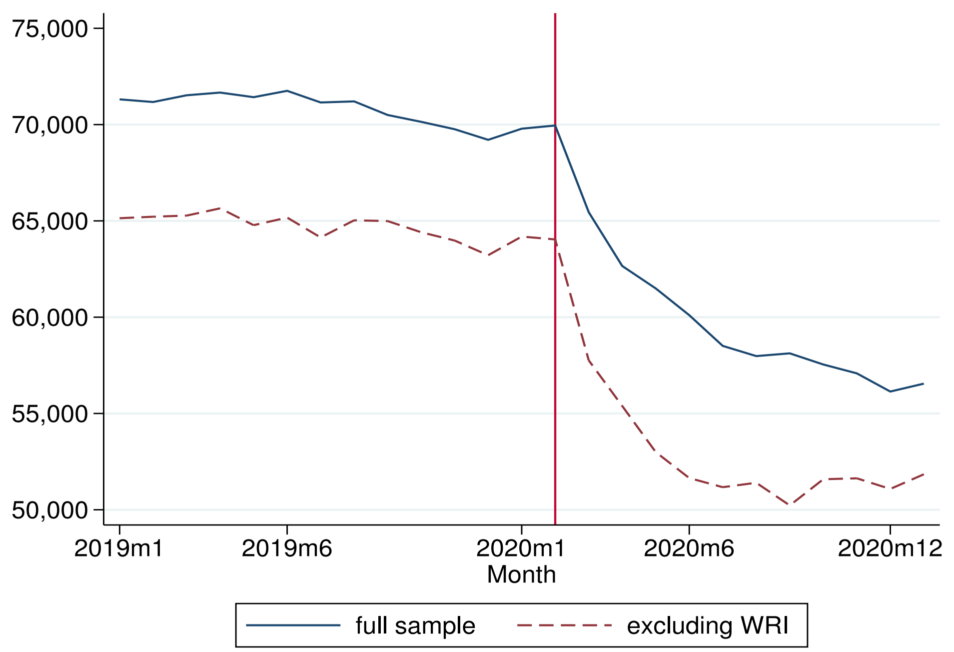

Figure 3: Sample size with and without whole record imputation (WRI). Source: LFS and authors' calculations. The blue plain line shows the total number of individuals in the full sample and the red dash line shows the number of individuals after excluding those for whom LFS information has been fully imputed. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

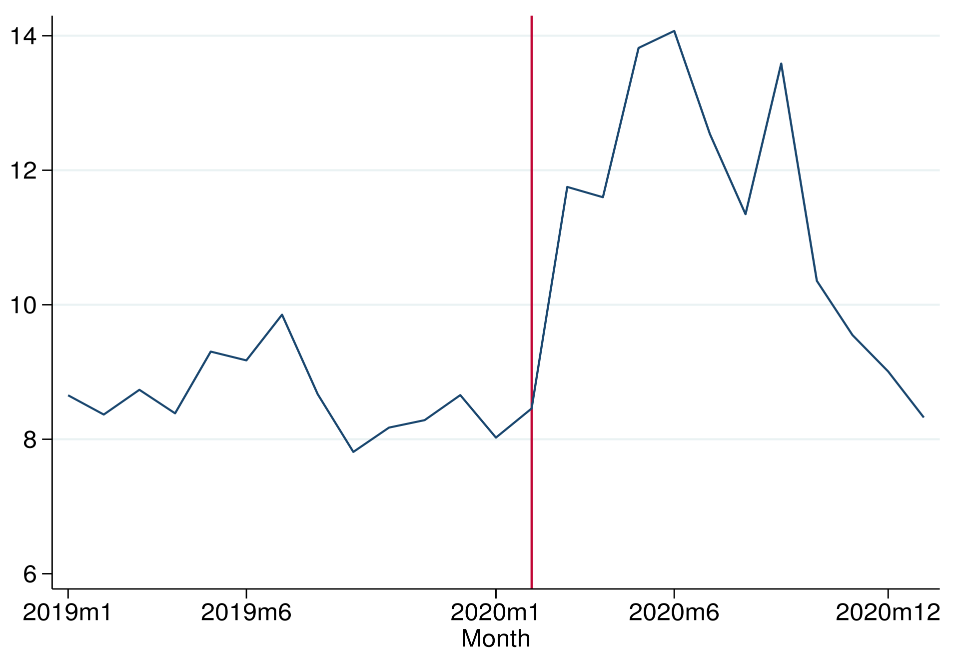

Figure 4: Whole record imputation (WRI) cases as a percentage of the full sample. Source: LFS and authors' calculations. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

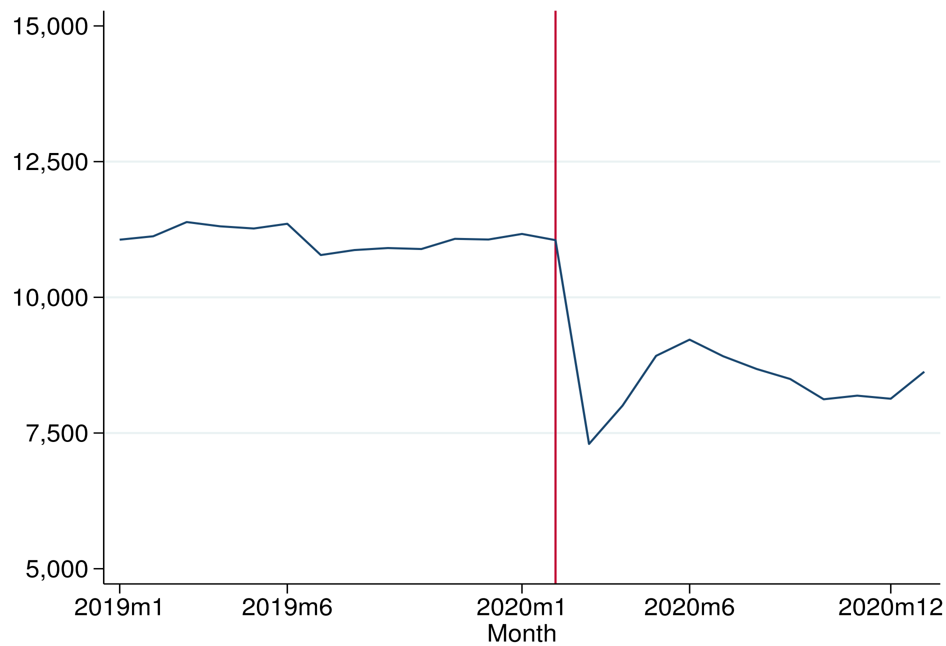

Figure 5: Sample size of incoming rotations. Source: LFS and authors' calculations. Number of individuals in the incoming rotation group in each month of the LFS. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

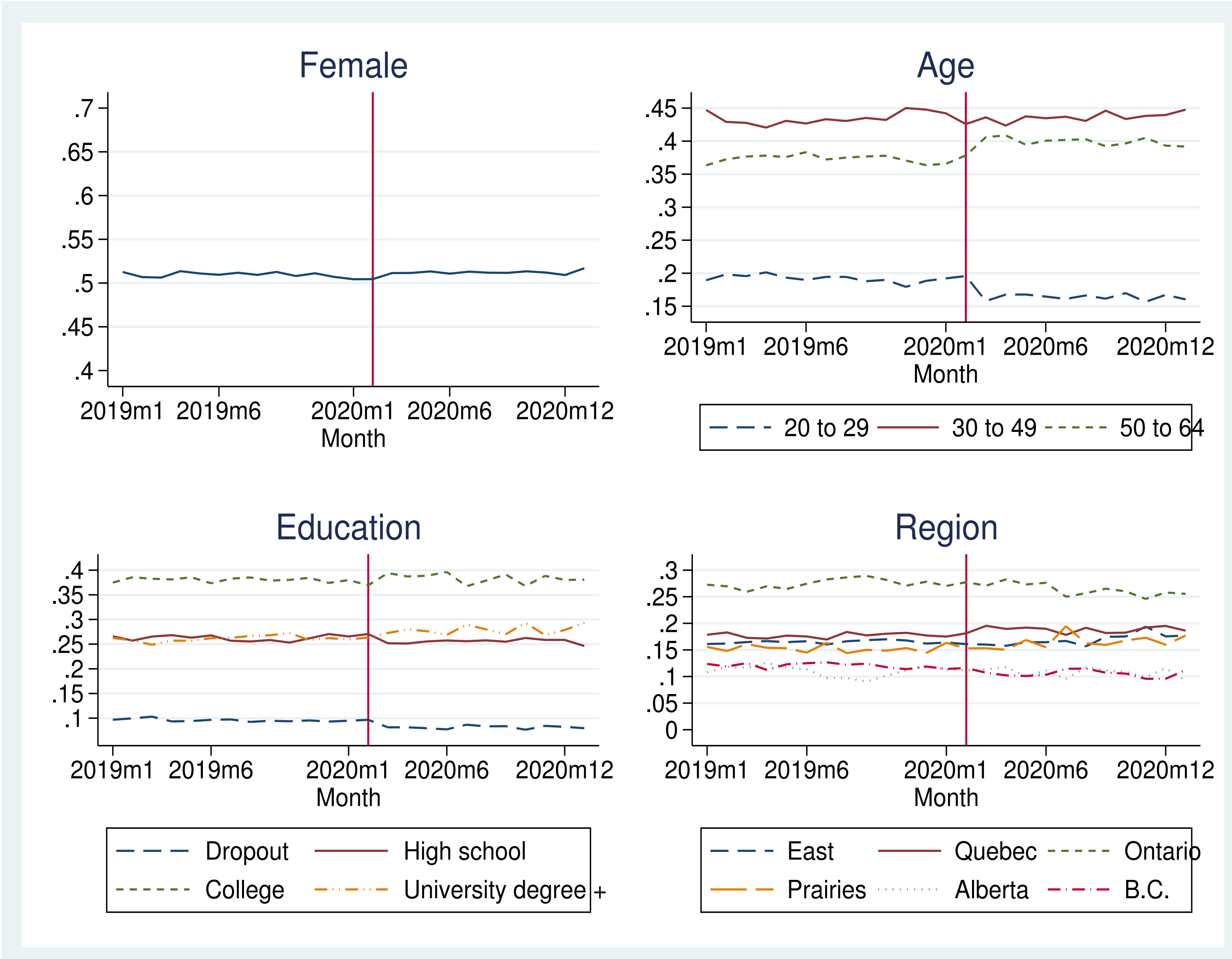

Figure 6: Proportions by subgroups, incoming rotation, unweighted. Source: LFS and authors' calculations. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

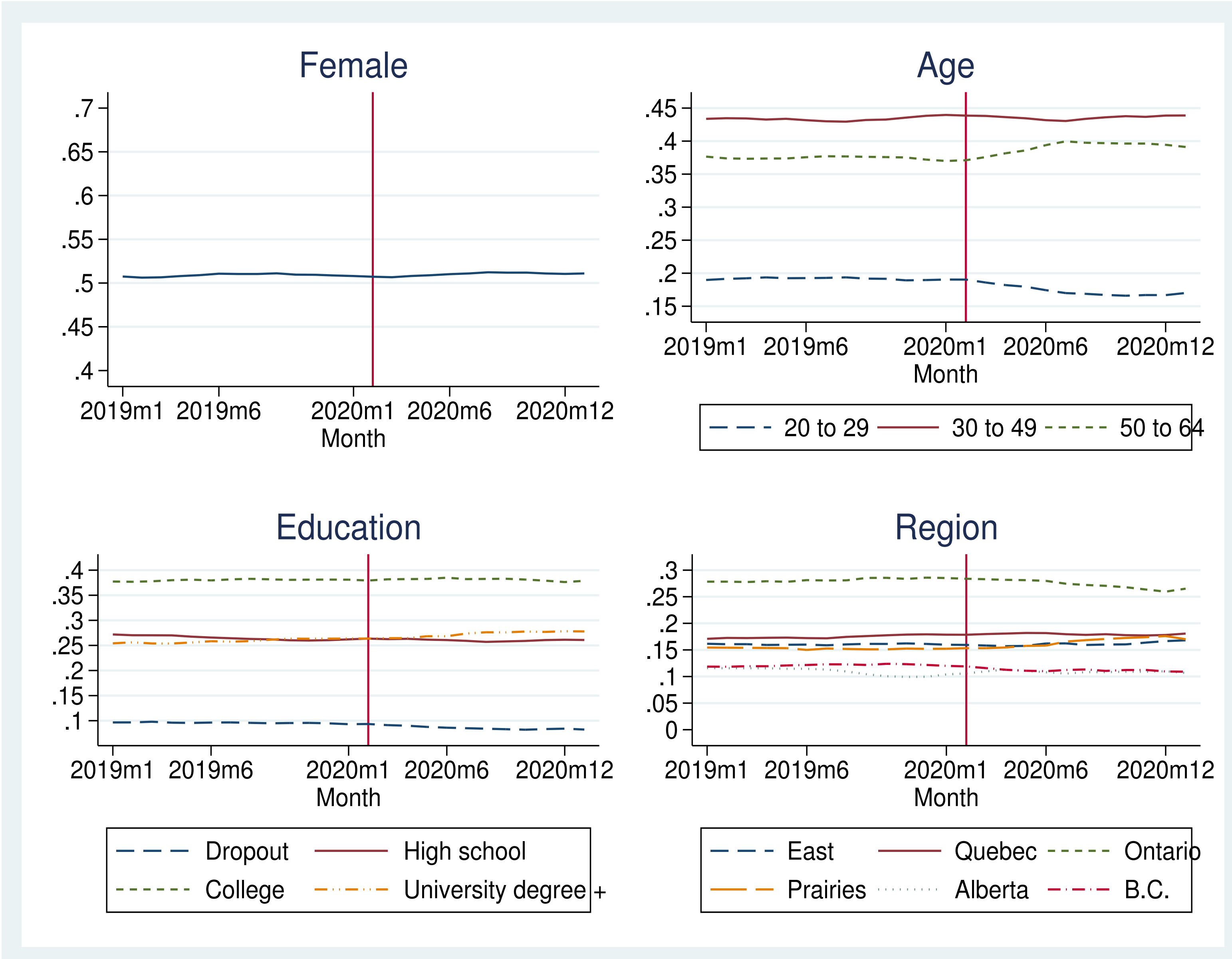

Figure 7: Proportions by subgroups, all rotations, unweighted. Source: LFS and authors' calculations. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

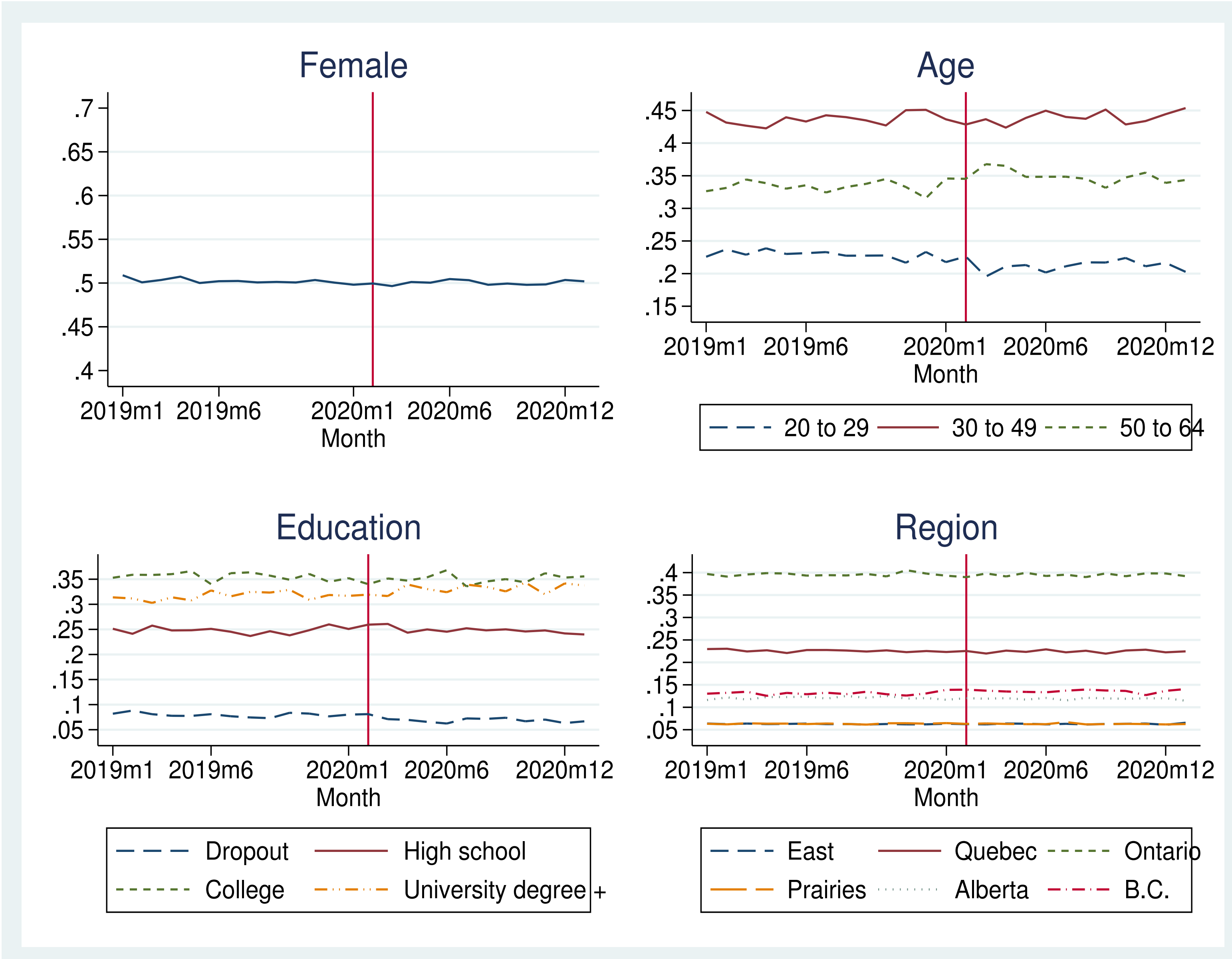

Figure 8: Proportions by subgroups, incoming rotation, weighted. Source: LFS and authors' calculations. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

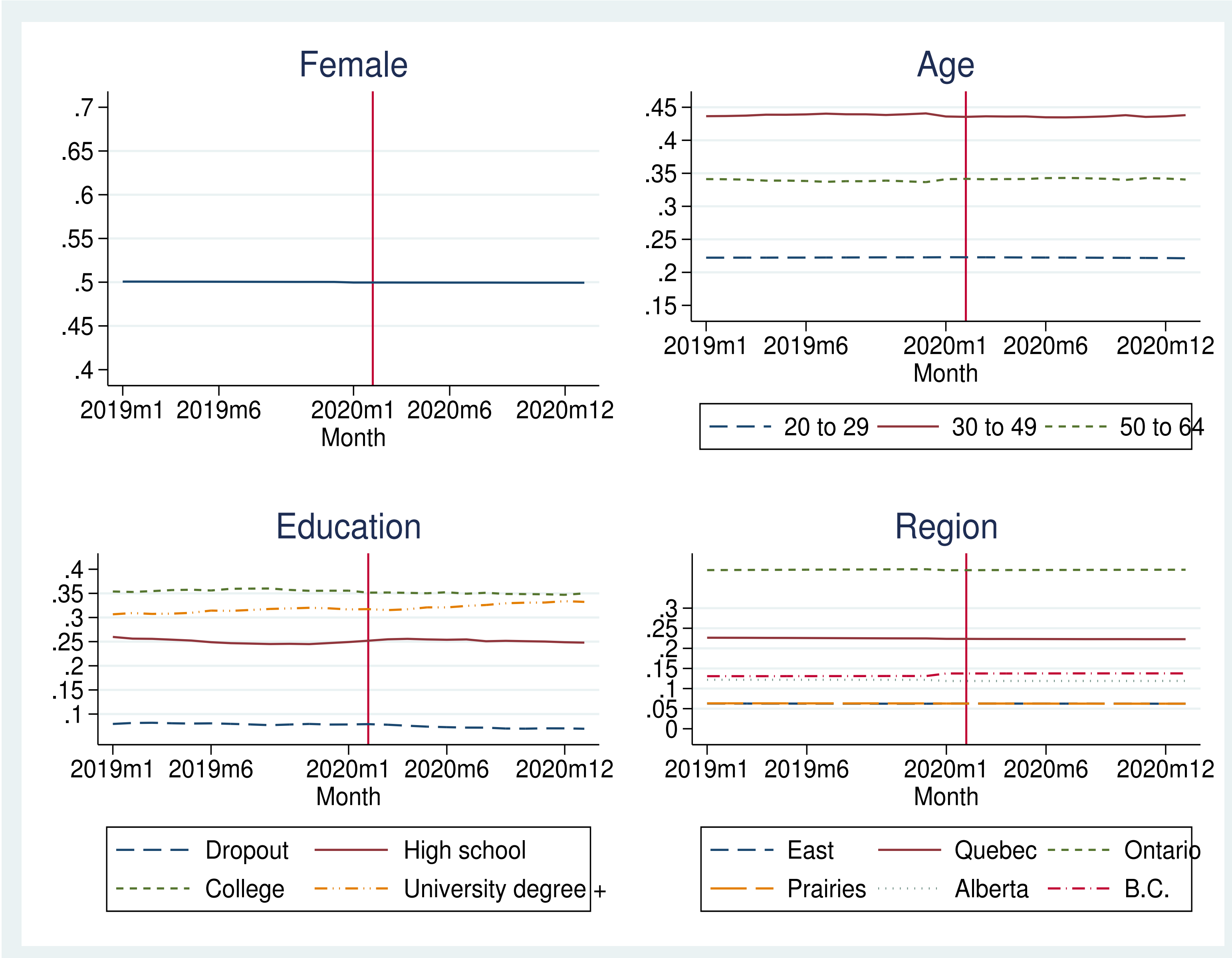

Figure 9: Proportions by subgroups, all rotations, weighted. Source: LFS and authors' calculations. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

Table 1: Non-response rates (%), by type

		Response	Birth non-response	Subsequent non-response	Birth non-response Incoming	Birth non-response Other
		(1)	(2)	(3)	(4)	(5)
2019	Jan	87.8	3.9	8.3	1.8	2.1
	Feb	88.0	4.0	8.0	1.7	2.3
	Mar	87.8	3.8	8.4	1.4	2.4
	Apr	88.4	3.5	8.1	1.4	2.1
	May	87.8	3.2	9.0	1.4	1.8
	Jun	88.2	2.9	8.9	1.3	1.6
	Jul	87.3	3.2	9.5	2.0	1.2
	Aug	88.9	2.7	8.4	1.8	0.9
	Sep	89.4	3.0	7.6	1.7	1.4
	Oct	89.3	2.8	7.9	1.6	1.2
	Nov	89.1	2.9	8.0	1.2	1.6
	Dec	88.4	3.2	8.4	1.2	2.0
2020	Jan	88.9	3.3	7.8	1.2	2.1
	Feb	88.1	3.8	8.1	1.5	2.3
	Mar	77.9	11.7	10.4	6.8	4.9
	Apr	76.5	13.5	10.0	5.6	7.9
	May	73.5	14.7	11.8	4.3	10.4
	Jun	71.8	16.4	11.8	3.8	12.6
	Jul	70.5	19.4	10.1	4.4	15.0
	Aug	70.5	20.5	9.0	4.8	15.7
	Sep	69.3	19.8	10.9	4.9	14.9
	Oct	70.0	21.9	8.1	5.6	16.3
	Nov	69.9	22.7	7.4	5.6	17.1
	Dec	69.2	23.9	6.9	5.7	18.3
2021	Jan	70.2	23.4	6.4	5.0	18.4

Notes: Monthly individual non-response rates (%) by type. Source: Statistics Canada and authors' own calculations (see online appendix for details about the calculation of rates in columns (2) to (5)). Sample for individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories.

Table 2: Probability of being interviewed in person, 2019, unweighted

	March			July		
	all	birth only	incoming only	all	birth only	incoming only
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.001	0.002	0.000	-0.003	-0.006	-0.012
	(0.002)	(0.008)	(0.008)	(0.002)	(0.009)	(0.009)
20 to 29	0.014***	0.058***	0.056***	0.017***	0.063***	0.067***
	(0.003)	(0.011)	(0.012)	(0.003)	(0.012)	(0.013)
50 to 64	-0.011***	-0.057***	-0.057***	-0.015***	-0.049***	-0.042***
	(0.002)	(0.019)	(0.010)	(0.002)	(0.012)	(0.010)
Dropout	0.027***	0.049***	0.054***	0.031***	0.080***	0.075***
	(0.004)	(0.015)	(0.016)	(0.004)	(0.016)	(0.017)
College	-0.005*	-0.019*	-0.020*	-0.001	0.005	0.001
	(0.002)	(0.010)	(0.011)	(0.003)	(0.011)	(0.012)
Bachelor degree +	-0.010***	-0.028**	-0.025**	-0.007***	-0.024**	-0.033**
	(0.003)	(0.011)	(0.012)	(0.003)	(0.012)	(0.013)
East	0.037***	0.196***	0.209***	0.021***	0.110***	0.109***
	(0.003)	(0.013)	(0.013)	(0.003)	(0.014)	(0.014)
Quebec	-0.012***	-0.020	-0.011	-0.021***	-0.093***	-0.119***
	(0.003)	(0.013)	(0.013)	(0.003)	(0.013)	(0.014)
Prairies	0.041***	0.223***	0.232***	0.023***	0.101***	0.110***
	(0.003)	(0.013)	(0.013)	(0.003)	(0.013)	(0.014)
Alberta	0.048***	0.297***	0.319***	0.013***	0.132***	0.132***
	(0.003)	(0.014)	(0.015)	(0.004)	(0.016)	(0.017)
British Columbia	0.038***	0.200***	0.215***	0.019****	0.086***	0.092***
	(0.003)	(0.014)	(0.015)	(0.003)	(0.014)	(0.016)
adj. R^2	0.011	0.074	0.081	0.006	0.034	0.040
N	71,523	12,844	11,387	71,149	12,702	10,779

Notes: Probability of being interviewed in person conditional on individual characteristics. The full sample (i.e. all) covers individuals aged 20 to 64, excluding full-time members of the armed forces and those living in the territories. The samples in columns (2) and (5) are further restricted to only include individuals who are birth interviews. The samples in columns (3) and (6) are further restricted to only include individuals who are in the first month of the six-month window (i.e. incoming rotation). The dependent variable is an indicator for whether the household the individuals belong to was interviewed in person. All regressions are unweighted and include a constant. Standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table 3: Probability of being observed in select months, 2019-2020, incoming rotation, unweighted

	Mar (1)	May (2)	Jul (3)	Sep (4)	Nov (5)
Post COVID-19	-0.122***	-0.069***	-0.091***	-0.096***	-0.123***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Interaction terms with post COVID-19 indicator:					
Female	0.003	0.001	-0.000	-0.001	-0.000
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
20 to 29	-0.047***	-0.033**	-0.044***	-0.036***	-0.021
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
50 to 64	0.013	0.006	0.016	0.002	0.025**
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Dropout	-0.035*	-0.023	-0.020	-0.013	-0.019
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
College	0.033***	0.026**	0.007	0.030**	0.023*
	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)
Bachelor degree +	0.035**	0.027**	0.028**	0.010	0.020
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
East	-0.008	0.001	0.051***	0.038**	0.068***
	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)
Quebec	0.028*	0.021	0.053***	0.035**	0.044***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Prairies	-0.020	0.019	0.076***	0.038**	0.055***
	(0.016)	(0.015)	(0.015)	(0.016)	(0.016)
Alberta	0.006	-0.022	0.051***	0.096***	0.018
	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)
British Columbia	-0.034**	-0.042***	0.019	-0.000	-0.002
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
adj. R^2	0.013	0.005	0.006	0.006	0.009
N	40,865	42,369	41,873	41,583	41,446

Notes: Probability of being observed in select months, conditional on individual characteristics. Sample for individuals aged 20 to 64 in the incoming rotation, excluding full-time members of the armed forces and those living in the territories, and observed in February and a select month of 2019 and 2020. The select month is identified at the top of each column. The dependent variable is an indicator for whether the individual is observed in the select month. All regressions are unweighted and include the sociodemographic characteristics on their own and a constant. Standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table 4: Probability of being observed in select months, 2019-2020, all rotations, unweighted

	Mar (1)	May (2)	Jul (3)	Sep (4)	Nov (5)
Post COVID-19	-0.019***	-0.035***	-0.054***	-0.059***	-0.064***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Interaction terms with post COVID-19 indicator:					
Female	-0.002	-0.002	-0.002	-0.002	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
20 to 29	-0.006	-0.014**	-0.028***	-0.032***	-0.028***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
50 to 64	0.005	0.014***	0.021***	0.019***	0.021***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Dropout	-0.011	-0.019***	-0.033***	-0.035***	-0.040***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
College	0.001	-0.003	-0.006	-0.006	-0.011**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Bachelor degree +	0.004	0.004	0.006	0.004	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
East	-0.002	0.001	0.018***	0.018***	0.027***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Quebec	0.003	0.007	0.015**	0.016***	0.015**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Prairies	0.001	0.011*	0.034***	0.051***	0.058***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Alberta	0.010	0.017**	0.017**	0.051***	0.070***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
British Columbia	-0.009	-0.020***	-0.013**	-0.007	-0.001
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
adj. R^2	0.000	0.001	0.003	0.003	0.004
N	278,100	274,042	270,784	269,746	267,969

Notes: Probability of being observed in select months, conditional on individual characteristics. Sample for individuals aged 20 to 64 in all rotations, excluding full-time members of the armed forces and those living in the territories, and observed in February and a select month of 2019 and 2020. The select month is identified at the top of each column. The dependent variable is an indicator for whether the individual is observed in the select month. All regressions are unweighted and include the sociodemographic characteristics on their own and a constant. Standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table 5: Probability of being interviewed in person, job characteristics, 2019, unweighted

	March			July		
	all	birth only	incoming only	all	birth only	incoming only
	(1)	(2)	(3)	(4)	(5)	(6)
Self-employed	-0.022***	-0.053***	-0.058***	-0.017***	-0.057***	-0.056***
	(0.004)	(0.019)	(0.020)	(0.005)	(0.020)	(0.021)
Employee	-0.026***	-0.062***	-0.058***	-0.024***	-0.052***	-0.052***
	(0.004)	(0.017)	(0.018)	(0.004)	(0.018)	(0.020)
Employee X earnings quartile 1	0.008***	0.021	0.027*	0.000	-0.008	-0.010
	(0.004)	(0.015)	(0.016)	(0.004)	(0.016)	(0.018)
Employee X earnings quartile 2	0.009***	0.032**	0.036**	0.003	-0.015	-0.012
	(0.003)	(0.014)	(0.015)	(0.004)	(0.015)	(0.016)
Employee X earnings quartile 4	-0.002	-0.013	-0.003	-0.004	-0.045***	-0.048***
	(0.003)	(0.015)	(0.016)	(0.004)	(0.016)	(0.017)
Employed X tenure (1 to 11 months)	0.022***	0.067***	0.062***	0.023***	0.077***	0.074***
	(0.003)	(0.014)	(0.014)	(0.003)	(0.014)	(0.015)
Employed X tenure (12 to 35 months)	0.015***	0.051***	0.055***	0.007**	0.029**	0.015
	(0.003)	(0.012)	(0.013)	(0.003)	(0.013)	(0.015)
Employed X telework	-0.008***	-0.020*	-0.020*	-0.008***	-0.015	-0.013
	(0.003)	(0.011)	(0.011)	(0.003)	(0.012)	(0.013)
Sociodemographic char. controls	Yes	Yes	Yes	Yes	Yes	Yes
Employed X industry controls	Yes	Yes	Yes	Yes	Yes	Yes
adj. R^2	71,523	12,844	11,387	71,149	12,702	10,779
N	0.014	0.082	0.089	0.008	0.040	0.045

Notes: Probability of being interviewed in person conditional on individual and job characteristics. The full sample (i.e. columns (1) and (4)) covers individuals aged 20 to 64 in all rotations, excluding full-time members of the armed forces and those living in the territories. The samples in columns (2) and (5) are further restricted to only include individuals who are birth interviews. The samples in columns (3) and (6) are further restricted to only include individuals who are in the first month of the six-month window (i.e. incoming rotation). The dependent variable is an indicator for whether the household the individuals belong to was interviewed in-person. All regressions are unweighted and include the sociodemographic characteristics on their own and a constant. Standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table 6: Probability of being observed in select months, 2019-2020, all rotations, weighted

	Mar (1)	May (2)	Jul (3)	Sep (4)	Nov (5)
Post COVID-19	0.004	0.005	0.008	0.008	0.006
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Interaction terms with post COVID-19 indicator:					
Female	0.000	-0.000	-0.000	-0.000	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
20 to 29	-0.001	-0.001	-0.001	-0.003	-0.002
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
50 to 64	-0.000	0.002	0.007	0.004	0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Dropout	-0.009	-0.019*	-0.030***	-0.028***	-0.033***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
College	-0.005	-0.011*	-0.018***	-0.019***	-0.015**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Bachelor degree +	-0.004	-0.004	-0.010	-0.009	-0.008
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
East	0.000	0.001	0.001	0.001	0.001
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Quebec	0.001	0.002	0.003	0.003	0.003
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Prairies	-0.000	0.000	0.000	0.000	0.000
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Alberta	0.000	0.001	0.001	0.001	0.001
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
British Columbia	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
adj. R^2	-0.000	-0.000	0.000	0.000	0.000
N	278,100	274,042	270,784	269,746	267,969

Notes: Probability of being observed in select months, conditional on individual characteristics. Sample for individuals aged 20 to 64 in all rotations, excluding full-time members of the armed forces and those living in the territories, and observed in February and a select month of 2019 and 2020. The select month is identified at the top of each column. The dependent variable is an indicator for whether the individual is observed in the select month. All regressions are weighted and include the sociodemographic characteristics on their own and a constant. Standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table 7: Pandemic impact across pre- and post-Covid entrants, 2019-2020, weighted

	(1) Mar	(2) Apr	(3) May	(4) Jun	(5) Jul	(6) Mar	(7) Apr	(8) May	(9) Jun	(10) Jul
Panel A: Probability of employment										
d^{2020}	-0.031***	-0.110***	-0.102***	-0.066***	-0.047***	-0.037***	-0.119***	-0.109***	-0.074***	-0.056***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.008)
d^r						-0.009	-0.003	-0.003	-0.007	-0.010*
						(0.006)	(0.005)	(0.004)	(0.004)	(0.006)
$d^{2020} \times d^r$						0.035***	0.026***	0.015**	0.012*	0.011
						(0.010)	(0.008)	(0.007)	(0.007)	(0.009)
N	136,975	134,322	132,917	131,860	129,659	136,975	134,322	132,917	131,860	129,659
Panel B: Probability of employment at work										
d^{2020}	-0.093***	-0.188***	-0.154***	-0.097***	-0.052***	-0.097***	-0.194***	-0.161***	-0.103***	-0.052***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.009)
d^r						-0.012*	-0.004	-0.004	-0.002	-0.005
						(0.007)	(0.005)	(0.005)	(0.005)	(0.006)
$d^{2020} \times d^r$						0.026**	0.017**	0.013*	0.009	-0.000
						(0.011)	(0.008)	(0.007)	(0.007)	(0.010)
N	136,975	134,322	132,917	131,860	129,659	136,975	134,322	132,917	131,860	129,659
Panel C: Probability of non-participation										
d^{2020}	0.016***	0.059***	0.042***	0.016***	0.007**	0.020***	0.068***	0.051***	0.022***	0.014**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.007)
d^r						0.010*	0.005	0.003	0.005	0.005
						(0.006)	(0.004)	(0.004)	(0.004)	(0.005)
$d^{2020} \times d^r$						-0.025***	-0.028***	-0.017***	-0.008	-0.009
						(0.009)	(0.007)	(0.006)	(0.006)	(0.008)
N	136,975	134,322	132,917	131,860	129,659	136,975	134,322	132,917	131,860	129,659

Notes: estimates of the impact of Covid-19 on select labour-market outcomes. Sample for individuals aged 20 to 64 in all rotations, excluding full-time members of the armed forces and those living in the territories, and observed in a specific calendar month, in 2019 and 2020. The relevant month is identified at the top of each column. Columns (1) to (5) report estimates of the model (4) with the restriction $\delta_1 = \delta_3 = 0$ i.e., estimates for the unconditional impact across all rotations. Columns (6) to (10) report estimates for the full model (4) i.e., it allows the impact to vary across the individuals assigned to pre- and post-Covid incoming rotations. All regressions are weighted and include a constant. Robust (Huber-White) standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table 8: Pandemic impact across pre- and post-Covid entrants, 2019-2020, 50-64 years-old, weighted

	(1) Mar	(2) Apr	(3) May	(4) Jun	(5) Jul	(6) Mar	(7) Apr	(8) May	(9) Jun	(10) Jul
Panel A: Probability of employment										
d^{2020}	-0.018***	-0.077***	-0.064***	-0.043***	-0.031***	-0.026***	-0.091***	-0.068***	-0.037***	-0.054***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.010)	(0.014)
d^r						-0.003	0.010	0.011	0.012	-0.006
						(0.011)	(0.008)	(0.008)	(0.008)	(0.010)
$d^{2019} \times d^r$						0.049***	0.040***	0.008	-0.009	0.028*
						(0.017)	(0.013)	(0.012)	(0.012)	(0.015)
N	51,307	50,694	50,425	50,631	50,209	51,307	50,694	50,425	50,631	50,209
Panel B: Probability of non-participation										
d^{2020}	0.012**	0.039***	0.028***	0.013**	0.006	0.018***	0.051***	0.035***	0.012	0.024*
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.013)
d^r						0.002	-0.009	-0.007	-0.005	0.005
						(0.010)	(0.008)	(0.007)	(0.008)	(0.010)
$d^{2019} \times d^r$						-0.031*	-0.033***	-0.013	0.001	-0.022
						(0.016)	(0.012)	(0.011)	(0.012)	(0.014)
N	51,307	50,694	50,425	50,631	50,209	51,307	50,694	50,425	50,631	50,209

Notes: estimates of the impact of Covid-19 on select labour-market outcomes. Sample for individuals aged 50 to 64 in all rotations, excluding full-time members of the armed forces and those living in the territories, and observed in a specific calendar month, in 2019 and 2020. The relevant month is identified at the top of each column. Columns (1) to (5) report estimates of the model (4) with the restriction $\delta_1 = \delta_3 = 0$ i.e., estimates for the unconditional impact across all rotation. Columns (6) to (10) report estimates for the full model (4) i.e., it allows the impact to vary across the individuals assigned to pre- and post-Covid incoming rotations. All regressions are weighted and include a constant. Robust (Huber-White) standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.