

CAN MATCHED
EMPLOYER-EMPLOYEE PANEL
SURVEY DATA ON TELEWORK
AND SELF-REPORTED
PRODUCTIVITY IDENTIFY
THE PRODUCTIVITY IMPACT OF
TELEWORK?

RENÉ MORISSETTE
ALI BÉJAOU

Les **cahiers de la série scientifique** visent à rendre accessibles les résultats des recherches effectuées par des chercheurs membres du CIRANO afin de susciter échanges et commentaires. Ces cahiers sont rédigés dans le style des publications scientifiques et n'engagent que leurs auteurs.

The purpose of the Working Papers is to disseminate the results of research conducted by CIRANO research members in order to solicit exchanges and comments. These reports are written in the style of scientific publications. The ideas and opinions expressed in these documents are solely those of the authors.

Le **CIRANO** est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du gouvernement du Québec, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Quebec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the government of Quebec, and grants and research mandates obtained by its research teams.

Les partenaires du CIRANO – CIRANO Partners

Partenaires Corporatifs - Corporate Partners

Autorité des marchés financiers
Banque de développement du
Canada
Banque du Canada
Banque Nationale du Canada
Bell Canada
BMO Groupe financier
La Caisse
Caisse de dépôt et placement du
Québec
Énergir
Hydro-Québec
Intact Corporation Financière
Mouvement Desjardins
Power Corporation du Canada
Pratt & Whitney Canada
VIA Rail Canada

Partenaires gouvernementaux - Governmental partners

Ministère des Finances du Québec
Ministère de l'Économie, de
l'Innovation et de l'Énergie
Innovation, Sciences et
Développement Économique
Canada
Ville de Montréal

Partenaires universitaires - University Partners

École de technologie supérieure
École nationale d'administration
publique de Montréal
HEC Montréal
Institut national de la recherche
scientifique
Polytechnique Montréal
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval
Université McGill

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web. CIRANO collaborates with many centres and university research chairs; list available on its website.

© April 2026. René Morissette, Ali Béjaoui. Tous droits réservés. *All rights reserved.* Reproduction partielle permise avec citation du document source, incluant la notice ©. *Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.*

Les idées et les opinions émises dans cette publication sont sous l'unique responsabilité des auteurs et ne représentent pas les positions du CIRANO ou de ses partenaires. *The observations and viewpoints expressed in this publication are the sole responsibility of the authors; they do not represent the positions of CIRANO or its partners.*

Can matched employer-employee panel survey data on telework and self-reported productivity identify the productivity impact of telework?

René Morissette^{}, Ali Béjaoui[†]*

Abstract

Measuring labour productivity is difficult. This is in part why recent studies on the productivity impact of telework focus on occupations where productivity is easily measured, such as call centers. One would think that matched employer-employee panel survey data on telework and self-reported productivity collected for stayers who experience an exogenous change in their number of telework days might provide an alternative. Since these data allow researchers to analyze changes in employees' self-reported productivity while controlling for firm fixed effects and ruling out reverse causality, they might identify the impact of telework on stayers' true productivity.

We show that even with such data, the impact of telework on productivity is not identified if telework affects self-reported productivity by affecting not only true productivity but also other factors—such as satisfaction with one's work arrangement—that likely depend on true productivity. Under these conditions, not controlling for these other factors leaves the other factors-to-self-reported productivity channel open but controlling for them leads to collider bias (Pearl and Mackenzie, 2018; Imbens, 2020), i.e. generates a spurious correlation between telework and self-reported productivity.

Résumé

Mesurer la productivité du travail est difficile. C'est en partie pour cette raison que les études récentes sur l'impact du télétravail sur la productivité se concentrent sur des professions où la productivité est facile à mesurer, comme les centres d'appels. On pourrait penser que des données d'enquête en panel appariées employeur-employé portant sur le télétravail et la productivité auto-déclarée, recueillies auprès de travailleurs restant dans la même entreprise mais connaissant un changement exogène du nombre de jours de télétravail, pourraient offrir une alternative. Comme ces données permettent aux chercheurs d'analyser les variations de la productivité auto-déclarée des employés tout en contrôlant les effets fixes liés à l'entreprise et en excluant la causalité inverse, elles pourraient permettre d'identifier l'impact du télétravail sur la productivité réelle des travailleurs restants.

Nous montrons que même avec ce type de données, l'impact du télétravail sur la productivité n'est pas identifiable si le télétravail influence la productivité auto-déclarée en affectant non seulement la productivité réelle, mais aussi d'autres facteurs tels que la satisfaction à l'égard de

^{*}Professeur associé, Université du Québec en Outaouais

[†]Professeur agrégé, Université du Québec en Outaouais, Fellow CIRANO

l'organisation du travail, qui dépendent probablement eux-mêmes de la productivité réelle. Dans ces conditions, ne pas contrôler ces autres facteurs laisse ouvert le canal reliant ces facteurs à la productivité auto-déclarée, mais les contrôler entraîne un biais de collision (collider bias) (Pearl et Mackenzie, 2018 ; Imbens, 2020), c'est-à-dire génère une corrélation artificielle entre le télétravail et la productivité auto-déclarée.

Mots-clés/Keywords : telework; work from home; productivity; collider bias/ télétravail; travail à domicile; productivité; biais de collision.

Pour citer ce document / To quote this document

Morissette, R. and Béjaoui, A. (2026). Can matched employer-employee panel survey data on telework and self-reported productivity identify the productivity impact of telework? (2026s-05, CIRANO). <https://doi.org/10.54932/HIJV8998>

I. Introduction

The increase in work from home triggered by the COVID-19 pandemic has raised an important question: what is the impact of telework on worker productivity? Answering this question is important for several reasons. It informs employers' decisions about whether to implement a return to the office, an issue that has attracted a lot of media attention over the last few years.³ It may also help explain why labour productivity stagnated in Canada from 2019 to 2023 or, alternatively, rule out telework as a factor underlying this stagnation.⁴

However, assessing the productivity impact of telework is challenging because measuring labour productivity is difficult, especially in service-producing jobs, which can often be done at home (Dingel and Neiman, 2020). This difficulty partly explains why many recent studies on the productivity impact of telework focus on occupations where individual productivity is easily measured, such as call centers (Dutcher, 2012; Bloom et al., 2015; Battiston et al., 2021; Gibbs et al., 2023; Emanuel and Harrington, 2024).⁵

It is assumed that matched employer-employee panel survey data that have repeated measures of telework and employees' self-reported productivity for stayers— employees who stay with the firm throughout the observation period—might provide an alternative.

As is well known, panel data may allow one to eliminate the time-invariant person-specific component that likely contaminates employees' self-reported productivity level. If employees report their productivity levels in *absolute* terms and if temporal changes in telework are measured precisely, first-differencing (or de-meaning) their responses will

³ <https://www.cbc.ca/news/canada/toronto/ontario-return-to-office-plan-9.7099817>;
<https://www.cbc.ca/news/politics/federal-government-return-to-office-9.7076314>.

⁴ <https://www150.statcan.gc.ca/n1/pub/11-631-x/11-631-x2024005-eng.htm>.

⁵ The paucity of datasets combining information on firms' value added and on the percentage of their workforce involved in telework is another explanation.

yield a cleaner estimate of the relationship between telework and self-reported productivity than one would obtain from cross-sectional data.⁶⁷

Armed with firm identifiers, researchers can then regress changes in employees' self-reported productivity on changes in telework intensity—the weekly number of days of telework—and on firm fixed effects for stayers.⁸ This will isolate the variation in self-reported productivity growth that occurs within firms for this group of employees. If true productivity growth—and therefore, self-reported productivity growth—varies across firms because of unobserved factors such as management quality and organizational culture, including firm fixed effects will control for these unobserved confounders in the short term, the period during which these confounders are time-invariant.

If stayers happen to experience an exogenous change in the number of days of telework, as was plausibly the case after the onset of the COVID-19 pandemic, concerns about reverse causality—for example, about a decrease in the productivity of a given employee leading to a decrease in her number of days of telework—can be ruled out. The resulting correlation between individual-level changes in telework intensity and in self-reported productivity is purged of variation due to person-specific fixed effects (associated with self-reported productivity *levels*), firm-specific fixed effects (associated with self-reported productivity *changes*) and reverse causality.

⁶ Although it is not a matched employer-employee dataset, the World Health Organization (WHO) Health and Performance Questionnaire (HPQ) is an example of dataset where workers report productivity levels in *absolute* terms: “On a scale from 0 to 10 where 0 is the worst job performance anyone could have at your job and 10 is the performance of a top worker, how would you rate your overall performance on the days you worked during the past 7 days.” See: <https://www.hcp.med.harvard.edu/hpq/info.php> .

⁷ If employees report productivity in *relative* terms—for example, by stating whether they are more productive at home than in the office—identifying the impact of telework on overall productivity is more challenging. The reason is that first-differencing employees' responses yields changes in relative productivity, which are not informative about changes in employees' *overall* productivity in a given week. For example, an increase in self-reported productivity at-home-relative-to-the-office could be entirely driven by a decrease in true productivity in the office—the denominator—that leads to a reduction in overall productivity.

⁸ Tracking leavers—employees who leave their initial firm during the observation period—in their new firm is prohibitively expensive in panel survey data and therefore is rarely, if ever, done.

Even after removing these sources of variation, changes in self-reported productivity will not necessarily capture the impact of telework on true productivity (henceforth, productivity). Employees may alter their responses about their productivity changes for a variety of reasons. For example, employees who become less satisfied with their work arrangement after a decrease in their number of days of telework may, for strategic reasons, report that their productivity fell, even though it remained unchanged. Alternatively, they may conceal a decrease in productivity, whenever such decrease occurs. Conversely, employees who become more productive and more satisfied with their work arrangement after an increase in their number of days of telework might overestimate their productivity gains. However, they are unlikely to conceal an increase in productivity, whenever such increase happens.

Such misreporting of productivity changes does not necessarily preclude the identification of the impact of telework. If changes in telework intensity trigger erroneous self-reported productivity changes by shifting an observable variable such as satisfaction with one's work arrangement, then controlling for this variable might identify the impact of telework on productivity. One might assume that, conditional on satisfaction with one's work arrangement, a one-unit increase in productivity will lead to a one-unit increase in self-reported productivity. We shall refer to this assumption as the one-to-one assumption. If this assumption is valid, the impact of telework on stayers' productivity might be identified even if their productivity is unobserved. Additionally, if the propensity to misreport productivity changes varies across worker characteristics and job characteristics, one may control for this possibility by conditioning results on those characteristics observed at the beginning of the observation period.

In sum, given their ability to sequentially control for employees' fixed effects and then employers' fixed effects for stayers while ruling out reverse causality under certain conditions, matched employer-employee panel survey data on telework and self-reported productivity might identify the impact of telework on stayers' productivity. The goal of this paper is to highlight the assumptions under which this impact is identified.

In Section II, we show that with the aforementioned data, the impact of telework on productivity is identified if: (i) the one-to-one assumption holds, (ii) telework affects self-reported productivity only by affecting productivity or by affecting productivity and other observable variables—such as satisfaction with one’s work arrangement—and (iii) these observable variables are unrelated to productivity or affect it. In contrast, the impact of telework on productivity is not identified if these other observable variables *are affected by* productivity. Under this alternative scenario, not controlling for these other variables leaves the other-variables-to-self-reported-productivity channel open but controlling for them leads to collider bias (Pearl and Mackenzie, 2018; Imbens, 2020), i.e. generates a spurious correlation between telework and self-reported productivity.

We also show that, under this alternative scenario, the effect of telework on self-reported productivity and on satisfaction with one’s work arrangement is identified. Along with the effect of telework on productivity, the effect of satisfaction on self-reported productivity is not identified because identification of this effect requires controlling for productivity, which is unobserved.

In Section III, we use simulated data to illustrate these findings in a simple fashion. We build a data-generating process (DGP) in which we know, by construction, the true parameters. In particular, we hard-code these parameters so that an increase in telework intensity causes an *increase* in self-reported productivity. Using the artificial sample resulting from this DGP, we run regressions that highlight which effects are identified. We show that regressing changes in self-reported productivity on changes in telework intensity while controlling for changes in satisfaction biases the estimated telework coefficient. Depending on the values of the true parameters underlying the DGP, the telework coefficient might be biased to such a large extent that telework appears to *decrease* self-reported productivity in these regressions.

In Section IV, we highlight the information set needed to capture the complex relationship between telework intensity, productivity and self-reported productivity. We start with the observation that employees’ overall productivity in a given week is an hours-weighted

average of productivity at home and in the office. Based on this consideration, we examine which self-reported productivity questions might best capture the impact of telework on productivity in panel data. Previous studies (e.g. Barrero, Bloom and Davis, 2021; Mehdi and Morissette, 2021; Béjaoui and St-Onge, 2025) often ask employees whether they are more, equally or less productive at home than in the office and therefore, measure self-reported productivity in relative terms, i.e. relative to productivity in the office. While these types of questions provide potentially useful information with cross-sectional data, they are not informative about how changes in telework intensity affect employees' overall productivity in a given week. The reason is simple: changes in telework intensity may simultaneously lead to, say, a decline in productivity in the office, an increase in productivity at home-relative-to-the-office and a decline in overall productivity. This implies that measuring employees' overall productivity in a given week requires asking *separate* questions about: a) employees' productivity at home; b) employees' productivity in the office; and c) the share of hours worked from home (or the office) that week. Combining these questions then allows researchers to assess how changes in telework intensity affect employees' overall productivity in a given week.

In Section V, we offer concluding remarks on our initial question: can matched employer-employee panel survey data on telework and self-reported productivity identify the productivity impact of telework?

II. Identifying assumptions

Assessing the impact of telework on labour productivity is an important task but measuring productivity is difficult, especially in service-producing jobs, which can often be done from home (Dingel and Neiman, 2020). This explains in part why recent research on the productivity impact of telework often analyzes jobs in settings where individual productivity is easily measured, such as call centers (Dutcher, 2012; Bloom et al., 2015; Battiston et al., 2021; Gibbs et al., 2023; Emanuel and Harrington, 2024).

Matched employer-employee panel survey data that have repeated measures of telework and employees' self-reported productivity might provide an alternative. They have several desirable features that might allow one to identify the productivity impact of telework for stayers.

By allowing researchers to compute first-differences or deviations from person-specific means, panel data may allow them to eliminate the time-invariant person-specific component that likely affects employees' self-reported productivity level.⁹ If employees report productivity levels in absolute terms—in terms of overall performance—and if temporal changes in telework are measured precisely, first-differencing (or de-meaning) their responses will yield a cleaner estimate of the relationship between telework and self-reported productivity than one would obtain from cross-sectional data

Researchers can regress changes in employees' self-reported productivity on changes in telework intensity and on firm identifiers for stayers. The resulting correlation between changes in telework intensity and changes in self-reported productivity is purged of person-specific fixed effects (associated with self-reported productivity *levels*) and firm-specific fixed effects (associated with self-reported productivity *changes*). If stayers experience an exogenous change in the number of days of telework, as is the case when senior managers alter telework arrangements through a formal scheme/policy covering a large number of employees rather than on an individual basis, concerns about reverse causality—for example, about a decrease in the productivity of a given employee leading her supervisor to reduce her number of days of telework—can be ruled out.

Within this empirical framework, if one puts forward the one-to-one assumption, i.e. assumes that a one-unit increase in productivity leads to a one-unit increase in self-reported productivity conditional on a set of controls, then the impact of telework on stayers' productivity might be identified even if their productivity is unobserved.

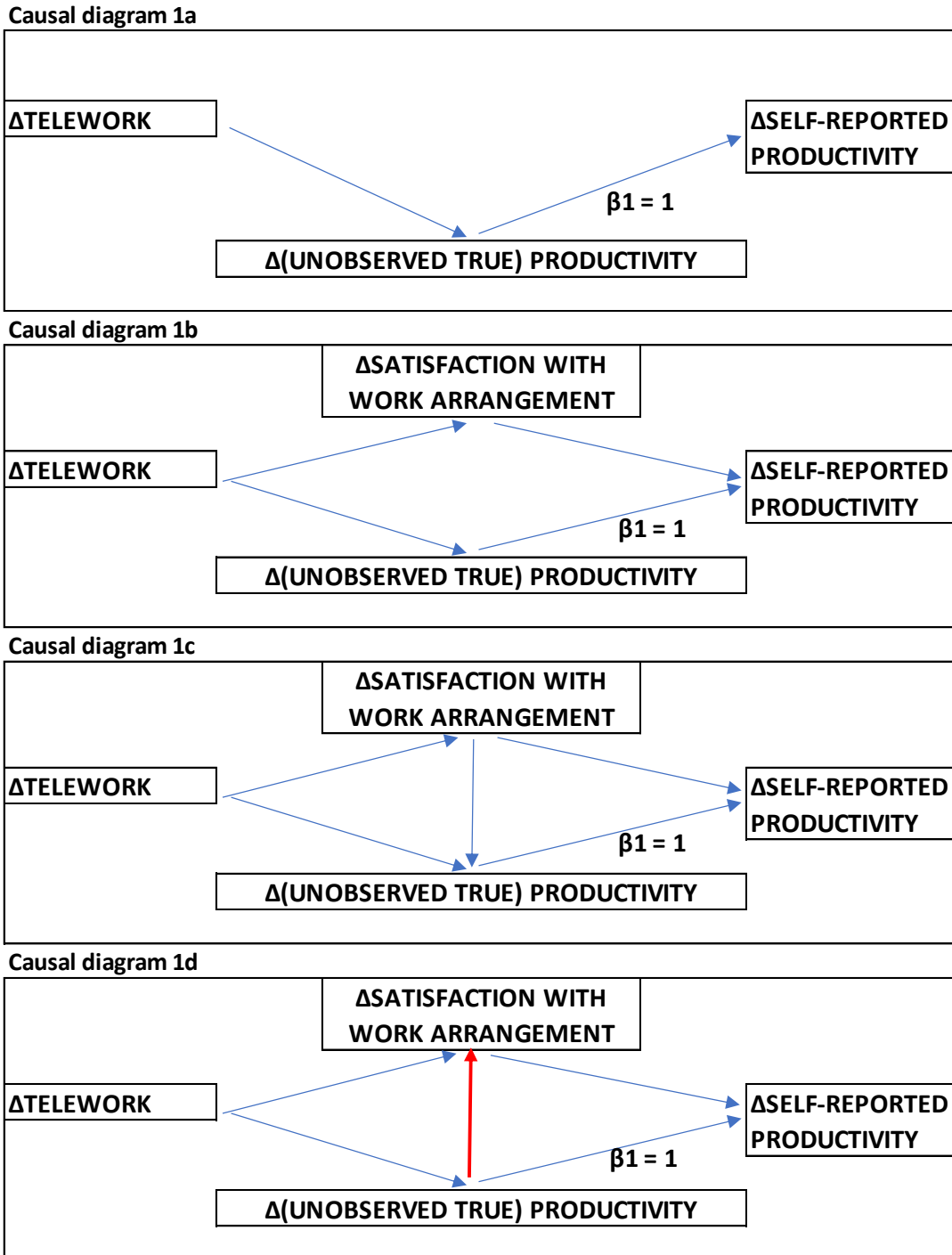
⁹This is feasible if this component enters the self-reported productivity function in an additive (rather than multiplicative) way.

The aforementioned data can identify the impact of telework on stayers' productivity under three sets of assumptions that are shown in the first three causal diagrams of Figure 1.

The first set of assumptions (A1) is fairly restrictive. It asserts that: (i) telework affects self-reported productivity only by affecting productivity and (ii) the one-to-one assumption holds ($\beta_1 = 1$). (Causal diagram 1a).

The second set of assumptions (A2) implies that: (i) telework affects self-reported productivity by affecting productivity and other observable factors—such as satisfaction with one's work arrangement—that (ii) are unrelated to productivity, and (iii) conditional on these observable factors, the one-to-one assumption holds ($\beta_1 = 1$) (Causal diagram 1b).

Figure 1: Telework, unobserved true productivity, and self-reported productivity---Identifying assumptions



Note: In Causal diagrams 1b and 1c, controlling for Δ SATISFACTION WITH WORK ARRANGEMENT allows one to identify the impact of TELEWORK on UNOBSERVED TRUE PRODUCTIVITY if $\beta_1=1$. In Causal diagram 1d, Δ SATISFACTION WITH WORK ARRANGEMENT is a **collider**: controlling for it will bias the estimated impact of TELEWORK on UNOBSERVED TRUE PRODUCTIVITY and on SELF-REPORTED PRODUCTIVITY (Pearl and Mackenzie, 2018; Imbens 2020). $\beta_1 = 1$: a one-unit increase in UNOBSERVED TRUE PRODUCTIVITY is assumed to lead to a one-unit increase in SELF-REPORTED PRODUCTIVITY.

The third set assumptions (A3) is similar to A2 but replaces condition (ii) by the following condition: satisfaction with one's work arrangement affects productivity but is not affected by it (Causal diagram 1c). If being more satisfied with one's telework arrangement leads an employee to put in more effort at work and therefore be more productive, then A3 will be a more plausible set of assumptions than A2.

Admittedly, A1 will be violated if telework affects self-reported productivity not only through productivity but also through other factors such as worker satisfaction with one's work arrangement.

For instance, employees who dislike commuting and who see their weekly number of days of telework drop from one year to the next will likely experience a decrease in satisfaction with their telework arrangement. As a result, they might strategically report a decrease in productivity during that period, even though their productivity did not fall. This could occur if they hope that their self-reported decrease in productivity might lead their employer to move back to the initial telework arrangement.¹⁰ Conversely, their counterparts who dislike telework because they miss social interactions will likely experience an increase in satisfaction and might strategically report a fictitious increase in productivity when telework intensity drops. For the same reasons, increases in telework intensity might lead to opposite behaviours for these two groups of employees.¹¹

This strategic behaviour of employees does not necessarily preclude the identification of the impact of telework on productivity. If satisfaction with one's work arrangement is unrelated to productivity—as A2 and Causal diagram 1b assume—or if satisfaction with one's work arrangement affects productivity but is not affected by it—as A3 and Causal diagram 1c assume—the effect of telework on productivity can be identified. This can be

¹⁰ Frustration about the new telework arrangement might also contaminate their responses about productivity changes, regardless of whether they hope that their employer will move back to the initial telework arrangement.

¹¹ Other factors (e.g. good or bad news related to one's personal life) that might lead employees to misreport productivity changes will be inconsequential if they are uncorrelated with temporal changes in telework intensity.

done by assuming that, conditional on satisfaction, the one-to-one assumption holds ($\beta_1 = 1$) and by regressing changes in self-reported productivity on telework changes while controlling for changes in satisfaction.

However, satisfaction with one's work arrangement may be affected by productivity (Causal diagram 1d). Workers who become more productive following a change in telework arrangement may report an increased satisfaction with this new work arrangement. If so, A2 and A3 no longer hold and the effect of telework on productivity is no longer identified. The reason is that in Causal diagram 1d, not controlling for satisfaction leaves the satisfaction-to-self-reported productivity channel open but controlling for it leads to *collider bias* (Pearl and Mackenzie, 2018; Imbens, 2020), i.e. generates a spurious correlation between telework and self-reported productivity. In other words, controlling for satisfaction in this context amounts to using a "bad control" (Angrist and Pischke, 2009).¹² We show in the next section that even if: a) telework increases self-reported productivity by increasing productivity and satisfaction and b) productivity increases satisfaction, controlling for satisfaction might generate a *negative* correlation between telework and self-reported productivity.

In sum, the impact of telework on productivity is not identified if telework affects self-reported productivity by affecting productivity and other factors, such as satisfaction with one's work arrangement, that *depend on* productivity (Causal diagram 1d).¹³

Two effects can be identified if the true data-generating process comes from Causal diagram 1d: 1) the effect of telework on satisfaction with one's work arrangement, which

¹² To understand this, consider that if satisfaction with one's work arrangement depends only on telework and productivity, which in turn depends on telework, asking to what extent a change in telework affects self-reported productivity, *holding satisfaction constant*, amounts to conducting an impossible thought experiment. The reason is that it is impossible to change telework *and keep satisfaction constant* because any change in telework will change satisfaction directly and indirectly (through the productivity channel). In this context, regressing changes in self-reported productivity on changes in telework and changes in satisfaction amounts to running a "bad" regression. Appendix 1 provides a more formal treatment of the issues involved, focusing on self-reported productivity changes.

¹³ This conclusion holds if one controls for worker and job characteristics at the beginning of the observation period (X), i.e. if we augment Causal diagrams 1a to 1d by having arrows pointing from X to Δ (unobserved true) productivity and Δ self-reported productivity.

is the sum of a direct effect and of an indirect effect mediated by productivity; 2) the effect of telework on self-reported productivity.

Along with the effect of telework on productivity, the effect of satisfaction on self-reported productivity is not identified in Causal diagram 1d because identification of this effect requires controlling for productivity, which is unobserved. For the same reason, the effect of satisfaction on self-reported productivity is not identified in Causal diagram 1c either.

In contrast, if satisfaction with one's work arrangement and productivity are unrelated, as is the case in Causal diagram 1b, then the impact of satisfaction on self-reported productivity can be recovered even though productivity is unobserved. This can be done using a five-step procedure.¹⁴ Hence, the existence of a causal link between satisfaction and self-reported productivity can be tested empirically only in Causal diagram 1b.

III. Simulated data

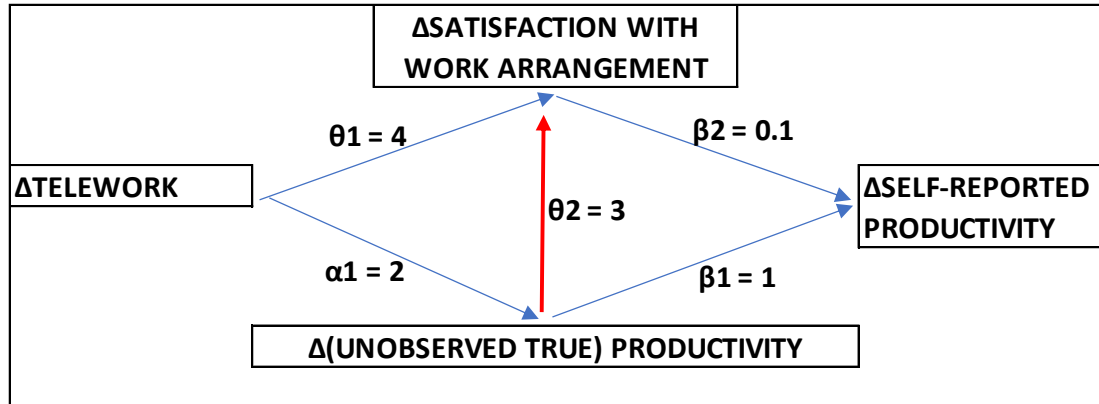
To illustrate which effects of telework are identified under Causal diagram 1d, we augment this diagram with a set of "true" parameters and assume that the data-generating process (DGP) is given by Figure 2. In Figure 2, a one-unit increase in telework intensity increases productivity by 2 units ($\alpha_1 = 2$), self-reported productivity by 3 units ($\alpha_1 \cdot \beta_1 + \theta_1 \cdot \beta_2 + \alpha_1 \cdot \theta_2 \cdot \beta_2 = 3$) and satisfaction with one's work arrangement by 10 units ($\theta_1 + \alpha_1 \cdot \theta_2 = 10$).

Using R, we then generate an artificial sample of 10,000 observations in which we hard-code these parameters in three equations—one equation for satisfaction, one for true

¹⁴ This impact can be recovered as follows: (i) regress changes in self-reported productivity on changes in telework intensity to obtain the total effect of telework intensity on self-reported productivity; (ii) regress changes in self-reported productivity on changes in telework intensity and changes in satisfaction to obtain the effect of telework intensity on self-reported productivity that is mediated by productivity; (iii) subtract (ii) from (i) to obtain the effect of telework intensity on self-reported productivity that is mediated by satisfaction; (iv) regress changes in satisfaction on changes in telework intensity to obtain the effect of telework intensity on satisfaction; (v) divide (iii) by (iv) to obtain the effect of satisfaction on self-reported productivity.

productivity and one for self-reported productivity—and add an error term $u \sim N(0,1)$ in each of these equations (see Appendix Table 1 for details).¹⁵

Figure 2: Simulating the data-generating process.



Effect of telework on self-reported productivity	$= \alpha_1 \cdot \beta_1 + \theta_1 \cdot \beta_2 + \alpha_1 \cdot \theta_2 \cdot \beta_2$
	$= 2 \cdot 1 + 4 \cdot 0.1 + 2 \cdot 3 \cdot 0.1$
	$= 3$
Effect of telework on true productivity	$= \alpha_1$
	$= 2$
Effect of telework on satisfaction	$= \theta_1 + \alpha_1 \cdot \theta_2$
	$= 4 + 2 \cdot 3$
	$= 10$

Note: In Figure 2, Δ SATISFACTION WITH WORK ARRANGEMENT is a collider: controlling for it will bias the estimated impact of TELEWORK on UNOBSERVED TRUE PRODUCTIVITY and on SELF-REPORTED PRODUCTIVITY (Pearl and Mackenzie, 2018; Imbens 2020). $\beta_1 = 1$: a one-unit increase in UNOBSERVED TRUE PRODUCTIVITY is assumed to lead to a one-unit increase in SELF-REPORTED PRODUCTIVITY.

Once this artificial sample is constructed, we run the following models:

MODEL 1: Regress Δ SELF_PR on Δ TELEWORK

MODEL 2: Regress Δ SELF_PR on Δ TELEWORK and Δ SATISFACTION

MODEL 3: Regress Δ PR on Δ TELEWORK

MODEL 4: Regress Δ PR on Δ TELEWORK and Δ SATISFACTION

¹⁵ Cunningham (2021, pp. 106-110) performs a similar exercise to show that controlling for occupation may contaminate the gender coefficient in an earnings equation.

where $\Delta\text{SELF_PR}$ and ΔPR denote changes in self-reported productivity and changes in productivity, respectively, and where $\Delta\text{TELEWORK}$ is the change in the weekly number of days of telework from year t to year $t+1$.¹⁶

With these models, we can test four implications of Figure 2.

- 1) MODEL 1 should yield a coefficient for $\Delta\text{TELEWORK}$ close to 3, i.e. the effect of telework intensity on self-reported productivity¹⁷;
- 2) MODEL 3 should yield a coefficient for $\Delta\text{TELEWORK}$ close to 2, i.e. the effect of telework intensity on productivity;
- 3) controlling for $\Delta\text{SATISFACTION}$ should bias the estimated coefficient for $\Delta\text{TELEWORK}$ in MODEL 2 and ...
- 4) ... in MODEL 4.¹⁸

Table 1 confirms all four patterns. The coefficient for $\Delta\text{TELEWORK}$ equals 3.008 in MODEL 1 and 2.002 in MODEL 3. Controlling for $\Delta\text{SATISFACTION}$ biases the estimated coefficient for $\Delta\text{TELEWORK}$. Under the set of parameters used in Chart 2, such control actually turns the initially positive coefficient for $\Delta\text{TELEWORK}$ into a *negative* one in MODEL 2 and MODEL 4.

Table 2 shows regression results obtained from an alternative DGP in which all parameters of Figure 2 are divided by 10, except β_1 and β_2 , which remain unchanged.¹⁹ In this new

¹⁶ We assume that $\Delta\text{TELEWORK}$ follows a uniform distribution that ranges from -5 days (moving from 5 days of telework to zero) to 5 days (moving from no day of telework to five days of telework). See Appendix Table 1.

¹⁷ The estimated coefficient will be close to 3 (not exactly 3) because of the error terms inserted in each of the three equations.

¹⁸ Since PR affects SATISFACTION but is not affected by it in Causal Diagram 1d, MODEL 4 is a purely descriptive regression that has no analytical content. We use it to show that in Causal diagram 1d, controlling for $\Delta\text{SATISFACTION}$ generates a spurious correlation between $\Delta\text{TELEWORK}$ and ΔPR . Note also that in an actual dataset, MODEL 3 and MODEL 4 are infeasible regressions because ΔPR is unobserved.

¹⁹ Appendix Table 2 provides the R codes used for this alternative DGP.

DGP, a one-unit increase in telework intensity increases self-reported productivity by 0.246 unit and productivity by 0.200 unit. Once again, MODEL 1 and MODEL 3 come very close to replicating these two parameters, with estimated coefficients for Δ TELEWORK that equal 0.245 and 0.198, respectively. In contrast, controlling for Δ SATISFACTION in MODEL 2 and MODEL 4 biases the estimated coefficient for Δ TELEWORK downwards, yielding estimated values that are now *positive* (0.063 and 0.065) but that are three to four times lower than the true parameters (0.246 and 0.200).

Table 1: Regressions illustrating collider bias for satisfaction in a productivity equation, under the data-generating process of Figure 2

DEPENDENT VARIABLE:	MODEL 1 Δ SELF_PR	MODEL 2 Δ SELF_PR	MODEL 3 Δ PR	MODEL 4 Δ PR
Δ TELEWORK	3.008*** (0.006)	-1.048*** (0.033)	2.002*** (0.004)	-1.001*** (0.010)
Δ SATISFACTION	-	0.405*** (0.003)	-	0.300*** (0.001)
CONSTANT	0.008 (0.017)	0.010 (0.010)	0.002 (0.010)	0.004 (0.003)
N	10,000	10,000	10,000	10,000

Note: Δ SELF_PR: changes in self-reported productivity; Δ PR: changes in true productivity; Δ SATISFACTION: changes in satisfaction with one's work arrangement; Δ TELEWORK: changes in the weekly number of days of telework. Data simulated from R codes shown in Appendix Table 1.

Table 2: Regressions illustrating collider bias for satisfaction in a productivity equation, under an alternative data-generating process

DEPENDENT VARIABLE:	MODEL 1 Δ SELF_PR	MODEL 2 Δ SELF_PR	MODEL 3 Δ PR	MODEL 4 Δ PR
Δ TELEWORK	0.245*** (0.005)	0.063*** (0.008)	0.198*** (0.003)	0.065*** (0.005)
Δ SATISFACTION	-	0.389*** (0.013)	-	0.285*** (0.009)
CONSTANT	0.018 (0.014)	0.022 (0.014)	-0.0004 (0.010)	0.003 (0.009)
N	10,000	10,000	10,000	10,000

Note: Δ SELF_PR: changes in self-reported productivity; Δ PR: changes in true productivity; Δ SATISFACTION: changes in satisfaction with one's work arrangement; Δ TELEWORK: changes in the weekly number of days of telework. Data simulated from R codes shown in Appendix Table 2.

Going back to the initial DGP outlined in Figure 2, we also run the following models:

MODEL 5: Regress Δ SATISFACTION on Δ TELEWORK

MODEL 6: Regress Δ SATISFACTION on Δ TELEWORK and Δ PR

MODEL 7: Regress Δ SELF_PR on Δ SATISFACTION

MODEL 8: Regress Δ SELF_PR on Δ SATISFACTION and Δ PR

to test four additional implications of Figure 2:

- 1) MODEL 5 should yield a coefficient for $\Delta\text{TELEWORK}$ close to 10, i.e. the (direct and indirect) effect of telework on satisfaction;
- 2) MODEL 6 should yield a coefficient for $\Delta\text{TELEWORK}$ close to 4, i.e. the direct effect of telework on satisfaction and a coefficient for ΔPR close to 3, the direct effect of productivity on satisfaction;
- 3) Since productivity and satisfaction are positively correlated in Chart 2, MODEL 7 should yield a coefficient for $\Delta\text{SATISFACTION}$ that is biased upwards. i.e. greater than 0.1, the effect of satisfaction on self-reported productivity;
- 4) MODEL 8 should yield a coefficient for $\Delta\text{SATISFACTION}$ close to 0.1, the effect of satisfaction on self-reported productivity and a coefficient for ΔPR close to 1, the effect of productivity on self-reported productivity.²⁰

Table 3 confirms all four patterns. The coefficient for $\Delta\text{TELEWORK}$ equals 10.006 in MODEL 5 and 3.984 in MODEL 6 while the coefficient for ΔPR equals 3.008 in MODEL 6. As expected, the coefficient for $\Delta\text{SATISFACTION}$ in MODEL 7 is, at 0.302, biased upwards: it is three times greater than 0.1. The coefficient for $\Delta\text{SATISFACTION}$ equals 0.099 in MODEL 8 and the coefficient for ΔPR equals 1.007 in that model.

²⁰ MODELS 6 and 8 are infeasible regressions in datasets where ΔPR is unobserved. They are used for illustrative purposes only.

Table 3: Regressions showing that the effect of telework on satisfaction is identified but effect of satisfaction on self-reported productivity is not identified---data-generating process of Figure 2

DEPENDENT VARIABLE:	MODEL 5 ΔSATISFACTION	MODEL 6 ΔSATISFACTION	MODEL 7 ΔSELF_PR	MODEL 8 ΔSELF_PR
ΔTELEWORK	10.006*** (0.011)	3.984*** (0.020)	-	-
ΔSATISFACTION	-	-	0.302*** (0.0004)	0.099*** (0.005)
ΔPR	-	3.008*** (0.010)	-	1.007*** (0.022)
CONSTANT	-0.006 (0.032)	-0.012 (0.010)	0.010 (0.011)	0.007 (0.010)
N	10,000	10,000	10,000	10,000

Note: ΔSELF_PR: changes in self-reported productivity; ΔPR: changes in true productivity; ΔSATISFACTION: changes in satisfaction with one's work arrangement; ΔTELEWORK: changes in the weekly number of days of telework. Data simulated from R codes shown in Appendix Table 1.

In sum, the simple simulations conducted in Tables 1 to 3 confirm that the effect of telework intensity on productivity is not identified under Causal diagram 1d. Assuming that telework increases self-reported productivity by increasing satisfaction and productivity and that productivity increases satisfaction, controlling for satisfaction yields an estimated impact of telework intensity on self-reported productivity that may be positive or negative, depending on the set of parameters selected. As a result, this estimated impact cannot be used to infer the effect of telework intensity on productivity, even if the one-to-one assumption holds ($\beta_1 = 1$). Likewise, the causal link between satisfaction and self-reported productivity cannot be identified under Causal diagram 1d. However, the effect of telework intensity on satisfaction with one's work arrangement is identified. The effect of telework intensity on self-reported productivity is also identified if one does not control for satisfaction.

IV. Information requirements

IV.1 Basic information requirements

The discussion above highlights the complexity of the relationship between telework intensity, productivity and self-reported productivity. One important question is: what information should be collected in matched employer-employee panel survey datasets to maximize the likelihood of identifying the productivity impact of telework?

Given its key role in Causal diagrams 1b to 1d, satisfaction with telework arrangement is an important piece of information. One important assumption underlying these three causal diagrams is that any increased (dis)satisfaction resulting from changes in telework intensity is *entirely* mediated by this variable. It is conceivable, however, that such (dis)satisfaction is mediated by other factors such as overall job satisfaction. If so, information on job satisfaction should also be collected.

If productivity growth and/or the propensity to misreport it vary across worker characteristics (gender, age, education, occupation, family status) and job characteristics (part-time status, union status), one will wish to control for this possibility by conditioning results on those characteristics observed at the beginning of the observation period.²¹ For this reason, collecting information on worker and job characteristics is crucial.

To identify stayers who experience an exogenous change in the number of days of telework, one should collect information on whether their telework arrangement is: a) part of a formal scheme/policy that covers a large number of employees in the organization; b) negotiated on an individual basis with their supervisor. Concerns about reverse causality—for example, about a decrease in the productivity of a given employee leading her supervisor to reduce her number of days of telework—can be minimized if one focuses on the former group.

²¹ Differences in industry of employment or firm size will be implicitly captured by the firm identifiers (firm fixed effects) used in the regression analyses.

The telework intensity measure should capture the weekly number of days working outside business premises (1, 2, 3, 4, 5, etc.) and should allow for the possibility of half days of telework (e.g. 2.5 days). To ensure that most of the work performed outside business premises is done at home, information on the amount of time (days or hours) worked from home should also be collected.

A key task is to select self-reported productivity questions that will best capture the impact of telework on productivity in panel data. The goal is to assess how changes in telework intensity affect employees' *overall* productivity in a given week.²²

One important consideration is that employees' *overall* productivity in a given week is an hours-weighted average of their productivity in the office and outside the office. Several previous studies (e.g. Barrero, Bloom and Davis, 2021; Mehdi and Morissette, 2021; Béjaoui and St-Onge, 2025) asked employees whether they are more, equally or less productive at home than in the office and therefore, measured self-reported productivity in *relative* terms, i.e. relative to productivity in the office. While these types of questions provide potentially useful information with cross-sectional data, they are not informative about how changes in telework intensity affect employees' *overall* productivity in a given week. The reason is that changes in telework intensity may simultaneously lead to, say, a decline in productivity in the office, an increase in productivity outside-the-office-relative-to-the-office and a decline in overall productivity.²³ This implies that measuring employees' overall productivity in a given week requires asking *separate* questions about: a) employees' productivity while teleworking; b) employees' productivity in the office; and c) the share of hours worked in and out of the office that week. Combining these questions

²² Admittedly, firms likely consider multiple factors (e.g. not only the impact of telework on productivity but also its impact on employee retention) when making decisions about telework intensity. This suggests that information on the quit rates of employees will also be useful.

²³ For example, a decrease in the number of days of telework might raise: a) logistical problems in the office if employees now have to find a work space (which is no longer theirs), adjust chairs and computer screens before starting to work; b) communication issues if colleagues of employees working in the office are not in the office during the same days or no longer sit nearby. These issues may reduce productivity in the office, leave productivity outside the office unchanged and reduce employees' overall productivity.

will then allow researchers to assess how changes in telework intensity affect employees' overall productivity in a given week.

What should be the exact wording of these questions? The World Health Organization (WHO) Health and Performance Questionnaire (HPQ) might provide some guidance here. It asks workers to report their productivity levels in *absolute* terms, using the following question: "On a scale from 0 to 10 where 0 is the worst job performance anyone could have at your job and 10 is the performance of a top worker, how would you rate your overall performance on the days you worked during the past 7 days."

This question could be adapted as follows to measure productivity while teleworking and productivity in the office:

Q1a: "On a scale from 0 to 10 where 0 is the worst job performance anyone could have at your job and 10 is the performance of a top worker, how would you rate:

- a) your overall performance in the office last week?
- b) your overall performance while teleworking last week?"

Alternatively, the following wording could also be used to focus on the amount of work performed per hour:

Q1b: "On a scale from 0 to 10 where 0 is the worst job performance anyone could have at your job and 10 is the performance of a top worker, how would you rate:

- a) the amount of work you could accomplish **per hour** in the office last week?
- b) The amount of work you could accomplish **per hour** while teleworking last week?"

Regardless of its wording, Q1 should be supplemented by a question that measures the number of hours worked last week in various work arrangements. One possible wording is:

Q2: "Last week, how many hours did you work:

- a) in the office?
- b) outside the office, i.e. while teleworking?

Taken together, Q1 and Q2 would allow researchers to compute employees' overall productivity in a given week as an hours-weighted average of their productivity in the office and outside the office.

IV.2 Other considerations

So far, the discussion has assumed that workers' unobserved heterogeneity affects the *levels* of self-reported productivity but not *changes* in self-reported productivity. If one relaxes this assumption, then person-specific fixed effects can be inserted in regressions that model changes in self-reported productivity.²⁴ However, such regressions are demanding in terms of data: they require having, for each worker, multiple first-differences, which in turn implies having at least 3 observations on self-reported productivity, telework intensity and other covariates for each employee.

Another assumption made throughout the study is that the effect of telework intensity on productivity is constant across workers. Recent research in econometrics relaxes this assumption and allows for heterogeneous treatment effects. When the treatment is continuous—as is the case with telework intensity—and workers' unobserved heterogeneity affects self-reported productivity *levels*, a difference-in-differences estimator developed by de Chaisemartin et al. (2025) can be used if, between consecutive periods, telework intensity changes for some employees but remains unchanged for others. Hence, it is possible to relax the assumption of a constant treatment effect if this condition is satisfied.

Lastly, whether one assumes that workers' unobserved heterogeneity affects only self-reported productivity levels or affects also self-reported productivity changes, a subtle measurement issue arises. While it is possible to quantify the degree to which a one-unit increase in telework intensity ($\Delta\text{TELEWORK}=1$) affects, on average, self-reported productivity ($\Delta\text{SELF_PR}$) measured in absolute terms, it is not possible to quantify the *average percentage change* in self-reported productivity (i.e. the average of $\Delta\text{SELF_PR} / \text{SELF_PR}$). The reason is that the two aforementioned assumptions about workers' unobserved heterogeneity imply that SELF_PR —the denominator—is not comparable across workers.

²⁴ In such regressions, firm-specific fixed effects are no longer needed or identified.

V. Concluding remarks

Do matched employer-employee panel survey data on telework and self-reported productivity provide an attractive option for researchers who wish to assess the impact of telework on productivity? The analyses performed in this study provide a nuanced answer to this question.

Assuming that changes in telework intensity affect employees' satisfaction with their telework arrangement (thereby ruling out Causal diagram 1a), two questions need to be tackled in Causal diagrams 1b to 1d:

- 1) does satisfaction with work arrangement affect self-reported productivity?
- 2) does (unobserved true) productivity affect satisfaction with work arrangement, as is the case in Causal diagram 1d?

We have shown that the first question can be answered empirically only under Causal diagram 1b, i.e. only if satisfaction and productivity are unrelated. In Causal diagrams 1c and 1d, where productivity and satisfaction are somehow related, the causal link between satisfaction and self-reported productivity is not identified.

Since productivity is unobserved, it is impossible to answer the second question empirically. Hence, researchers have to make two untestable assumptions in order to proceed with their empirical analyses. First, they have to assume that employees who become more productive *because* of a change in telework intensity will not report an increased satisfaction with this new telework arrangement. Second, they have to assume that conditional on observables, the one-to-one assumption holds. Some researchers will view one or both of these assumptions as implausible. They will conclude that the productivity impact of telework is not identified in matched employer-employee panel survey data. Other researchers might feel comfortable defending these two assumptions. They will proceed with their empirical analyses, using either Causal diagram 1b or 1c.

REFERENCES

- Barrero, J. M., N. Bloom, and S. J. Davis. 2021. 'Why working from home will stick.' NBER Working Paper 28731.
- Battiston, D., J. Blanes i Vidal, and T. Kirchmaier. 2021. 'Face-to-face communication in organizations', *Review of Economic Studies*, vol. 88(2), pp. 574–609.
- Béjaoui, A. and S. St-Onge. 2025. 'Les effets du travail à distance sur les attitudes du personnel au travail et le rôle modérateur du soutien organisationnel : résultats d'une recherche longitudinale', (2025RP-28, Projects Reports, CIRANO.) <https://doi.org/10.54932/SAEH4396>
- Bloom, N., J. Liang, J. Roberts, and Z.J. Ying. 2015. 'Does working from home work? Evidence from a Chinese experiment', *Quarterly Journal of Economics*, vol. 130(1), pp. 165–218.
- Cunningham, S. 2021. *Causal inference: the mixtape*. Yale University Press.
- de Chaisemartin, C., X. D'Haultfoeuille, F. Pasquier, D. Sow and G. Vazquez-Bare. 2025. 'Difference in differences for continuous treatments and instruments with stayers'.
- Dingel, J.I., and B. Neiman. 2020. 'How many jobs can be done at home?' *Journal of Public Economics*, vol. 189, September.
- Dutcher, E.G. 2012. 'The effects of telecommuting on productivity: An experimental examination. The role of dull and creative tasks', *Journal of Economic Behavior and Organization*, vol. 84(1), pp. 355–63.
- Emanuel, N. and E. Harrington. 2024. 'Working remotely or remotely working? Selection, treatment, and the market for remote work', *American Economic Journal: Applied Economics*, vol. 16(4), pp. 528–59.
- Gibbs, M., F. Mengel, and C. Siemroth. 2023. 'Work from home and productivity: Evidence from personnel and analytics data on information technology professionals', *Journal of Political Economy Microeconomics*, vol. 1(1), pp. 7–41.
- Imbens, G. 2020. 'Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics', *Journal of Economic Literature*, Vol. 58(4), pp. 1129-1179.
- Mehdi, T., and R. Morissette. 2021. 'Working from home: Productivity and preferences', StatCan COVID-19: Data to Insights for a Better Canada, no. 12. Statistics Canada Catalogue no. 45280001. Ottawa: Statistics Canada.
- Pearl, J., and D. Mackenzie. 2018. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Appendix 1: Causal diagrams 1c and 1d

Consider a general model of self-reported productivity changes:

$$\Delta SELFPR_i = f(\Delta PR_i, DEV\Delta PR_i) \quad (1)$$

where $\Delta SELFPR_i$, ΔPR_i , and $DEV\Delta PR_i$ denote self-reported productivity changes, (true) productivity changes and deviations from true productivity changes for employee i , respectively.

Assume that productivity changes and deviations from true productivity changes can be modelled as follows:

$$\Delta PR_i = g(X_{it0}, \theta_j, \Delta TEL_i, \Delta SAT_i) \quad (2)$$

$$\Delta DEVPR_i = h(X_{it0}, \theta_j, \Delta SAT_i) \quad (3)$$

where X_{it0} , ΔTEL_i , ΔSAT_i and θ_j denote worker and job characteristics measured at the beginning of the observation period, changes in telework intensity, changes in satisfaction with one's work arrangement and firm fixed effects, respectively.

Further assume that:

$$\Delta SAT_i = i(\Delta TEL_i) \quad (4)$$

Insert (2) to (4) into (1):

$$\Delta SELFPR_i = f\left(g\left(X_{it0}, \theta_j, \Delta TEL_i, \Delta SAT_i(\Delta TEL_i)\right), h\left(X_{it0}, \theta_j, \Delta SAT_i(\Delta TEL_i)\right)\right) \quad (1.1)$$

Abstracting from X_{it0} and θ_j , equation (1.1) corresponds to Causal diagram 1c. It shows that controlling for ΔSAT_i removes variation in ΔTEL_i that is related to ΔSAT_i but does not shut down variation in ΔTEL_i that directly affects ΔPR_i and therefore, $\Delta SELFPR_i$.

Now assume that:

$$\Delta PR_i = g(X_{it0}, \theta_j, \Delta TEL_i) \quad (5)$$

$$\Delta DEVPR_i = h(X_{it0}, \theta_j, \Delta SAT_i) \quad (6)$$

$$\Delta SAT_i = i(\Delta TEL_i, \Delta PR_i) \quad (7)$$

Insert (5) to (7) into equation (1):

$$\Delta SELFPR_i = f\left(g\left(X_{it0}, \theta_j, \Delta TEL_i\right), h\left(X_{it0}, \theta_j, \Delta SAT_i(\Delta TEL_i, \Delta PR_i)\right)\right) \quad (1.2)$$

Abstracting from X_{it0} and θ_j , equation (1.2) corresponds to Causal diagram 1d. It shows that controlling for ΔSAT_i implies either: 1) holding constant both ΔTEL_i and ΔPR_i or 2) changing ΔTEL_i and ΔPR_i in a *selective* way that maintains ΔSAT_i unchanged. The first option shuts down all sources of variation in ΔTEL_i while the second option generates a spurious correlation between ΔTEL_i and ΔPR_i in equation (5) that will spillover into $\Delta SELFPR_i$ in equation (1.2)

Appendix Table 1: R codes for simulated data

```
library(tidyverse)
library(stargazer)

tb <- tibble(
  d_telework = runif(10000, min=-5, max=5),
  d_pr = 0 + 2*d_telework + rnorm(10000),
  d_sat = 0 + 3*d_pr + 4*d_telework + rnorm(10000),
  d_selfpr = 0 + 1*d_pr + 0.1*d_sat + rnorm(10000)
)

lm_1 <- lm(d_selfpr ~ d_telework,tb)
lm_2 <- lm(d_selfpr ~ d_telework + d_sat,tb)
lm_3 <- lm(d_pr ~ d_telework,tb)
lm_4 <- lm(d_pr ~ d_telework + d_sat,tb)
lm_5 <- lm(d_sat ~ d_telework,tb)
lm_6 <- lm(d_sat ~ d_telework + d_pr,tb)
lm_7 <- lm(d_selfpr ~ d_sat,tb)
lm_8 <- lm(d_selfpr ~ d_sat + d_pr,tb)

stargazer(lm_1,lm_2,lm_3,lm_4,lm_5,lm_6,lm_7,lm_8,type="text",
          column.labels = c("MODEL 1",
                            "MODEL 2",
                            "MODEL 3",
                            "MODEL 4",
                            "MODEL 5",
                            "MODEL 6",
                            "MODEL 7",
                            "MODEL 8"))
```

Note: Codes run with RStudio Version 2025.09.2+418.

Appendix Table 2: R codes for simulated data

```
library(tidyverse)
library(stargazer)

tb <- tibble(
  d_telework = runif(10000, min=-5, max=5),
  d_pr = 0 + 0.2*d_telework + rnorm(10000),
  d_sat = 0 + 0.3*d_pr + 0.4*d_telework + rnorm(10000),
  d_selfpr = 0 + 1*d_pr + 0.1*d_sat + rnorm(10000)
)

lm_1 <- lm(d_selfpr ~ d_telework,tb)
lm_2 <- lm(d_selfpr ~ d_telework + d_sat,tb)
lm_3 <- lm(d_pr ~ d_telework,tb)
lm_4 <- lm(d_pr ~ d_telework + d_sat,tb)

stargazer(lm_1,lm_2,lm_3,lm_4,type="text",
          column.labels = c("MODEL 1",
                            "MODEL 2",
                            "MODEL 3",
                            "MODEL 4"))
```

Note: Codes run with RStudio Version 2025.09.2+418.