

Productivity and Work From Home: the Debate Continues



Some opinionated comments on the popular press, the academy, and the ongoing debate about productivity and working from home.

by Michael Colacino, Haiku Partners / October 2023

Productivity and Work From Home: the Debate Continues

The discussion over productivity and work from home (WFH) is ongoing, as it has been since the onset of COVID and the forced relocation of office workers to home. Over time it has become an increasingly polarized issue, and extremely popular for both academics and the popular press.

One thing is clear from reading the business press, and that is that many of the business writers are wrong! When they assert things like “it’s been proved that productivity has been improved by working from home” it’s my contention that it isn’t even close to being proved by the research we have to date, and a careful reading of the research and evaluation of corporate actions shows certain patterns both in the data and in the decisions being made by organizations. Those patterns indicate a lot of things but do not show that productivity increases when people work from home. Why are so many writers so certain of this “fact”?

Return to office mandates are ridiculous, unnecessary and, in some instances, even cruel

Julie Bort Sep 5, 2023, 10:36 AM EDT



Nap pods at Google's Mountain View Googleplex in 2015. Brooks Kraft LLC/Corbis via Getty Images

- I have been a full-time remote worker for nearly all of my career.
- Return-to-office mandates are not really about productivity or better meetings.

from *Business Insider* (Bort 2023)

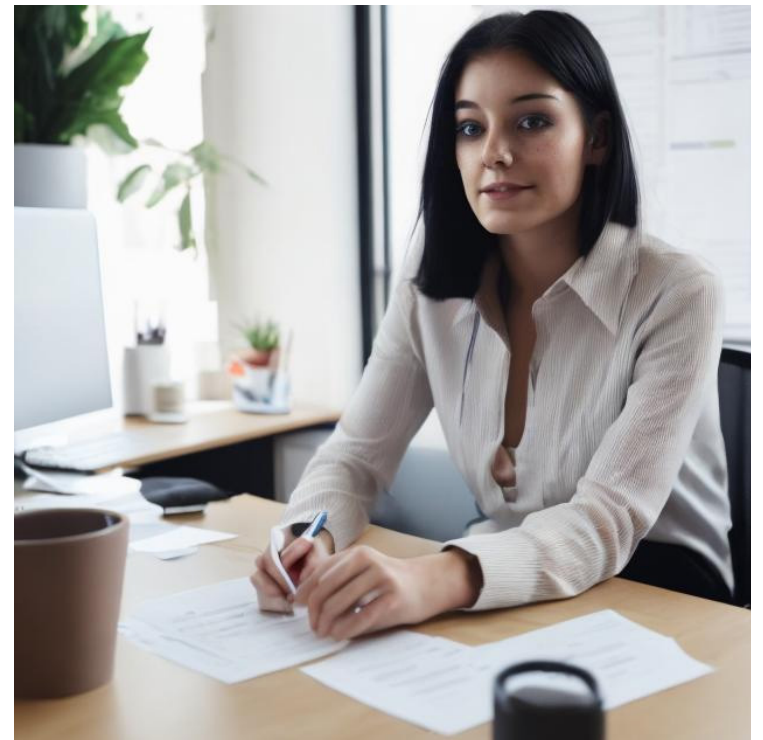
This paper will try to explain this, beginning with the academic world, and then reflecting on how that is being presented or distorted in the business and popular press. In the world of the academy and in the corporate world, the data being gathered (and the follow-on analysis) seems to fall in three broad categories.

Three types of data

1. Self Assessment

The first category could be labeled “self-assessment.” These studies are based in whole or in part on surveys, some quite large in sample size, which ask a variety of questions to employees and management. Two of the common questions are “Do you enjoy or prefer working from home some (all) of the time?” and “Do you consider yourself more productive when working from home?”

Examples of this type of research include Bloom and Barrero and their “Survey of Working Arrangements and Attitudes.” These survey-based studies have the advantage of being relatively easy and inexpensive to accomplish on a large scale, and so they were first to be done and analyzed. Their disadvantage is that they lack some of the traits of a scientific study, such as control groups and double-blind protocols. Studies that fit in this category include (Aksoy *et al.*, 2022), (Angelici and Profeta, 2020), (Barrero, Bloom and Davis, 2021), (Gavoille and Hazans, 2022), (Smite *et al.*, 2022), (Stropoli, 2021), (Kitagawa *et al.*, 2021) and (Morikawa, 2022).



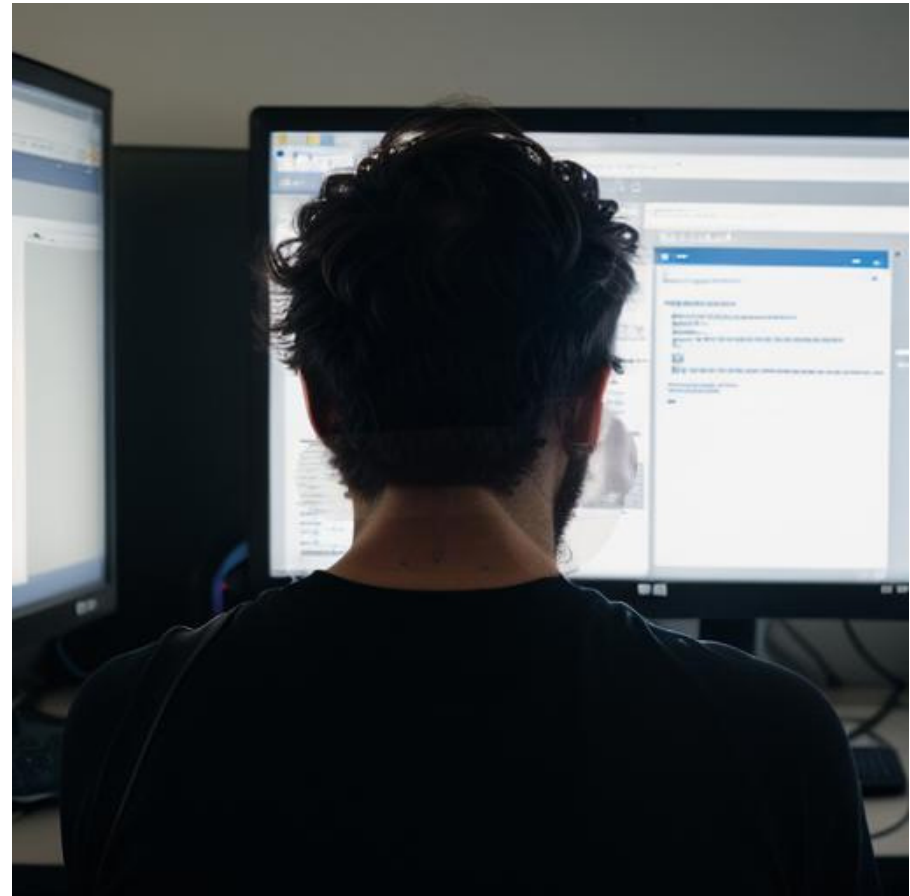
Kitagawa *et al* and Morikawa (both studies of Japanese workers) are the outliers in this group, and unlike many of the other self-assessment studies reflect a perceived reduction in productivity.

Smite *et al* perform a study of thirteen studies, examining perceived productivity of software engineers (a group that one might think would lean natural-toward WFH), finding that about 50% of the respondents perceived productivity to be higher, and around the same percentage found it to be lower. They come to their analysis from the perspective that remote software development is less productive and express surprise at the opposite result in some of the data:

“The decreased productivity when moving from collocated to fully distributed setup is not that surprising. Distributed work is infamous for being considerably less productive than collocated work due to impaired teamwork (Herbsleb and Mockus, 2003). What is in fact surprising, is that some WFH developers perceive being more productive.” (Smite et al., 2022, p. 11) quoting (Herbsleb and Mockus, 2003)

Angelici *et al* rely heavily on self-reported and supervisor-reported data, but also bring in two measures of objective data¹.

¹ The methodology is explained in Appendix A. Their assessment of the participants' days of leave requested is straightforward; their analysis of 'productivity' is less convincing.



One of the interesting ways in which these self-assessment studies investigate their survey participants is to ask them how much more productive they are working from home compared with what they *expected*. For example, roughly sixty percent of the respondents reported being more productive than they expected, versus fourteen percent who reported the opposite in (Barrero, Bloom and Davis, 2021, p. 17). This idea of a “productivity surprise” is also found in (Aksoy *et al.*, 2022, p. 3). Whether being surprised at the difference between employee productivity expectations and post hoc employee perception is meaningful data is not clear, and fits inside the broader topic of self-assessment, as discussed below.

2. Induction

The second category, could be called “induction studies.” These attempt to find some measurable quantity, such as lines of code, comments on Github, or email trajectories, that could be compared between two groups which differ in their level of working from home, either currently or in comparison with pre-COVID levels. These studies attempt to infer what’s causing the difference. This methodology is more challenging to implement since controlled experiments, such as Nicholas Bloom’s 2013 C-Trip study (Bloom *et al.*, 2014), are logistically much harder to perform. Bloom’s study is so often cited, in part, because (in the remote work research space) it was a unique opportunity to have a control and treatment group of significant size over a relatively long time-frame.

While this is a groundbreaking study, and has established Bloom as a primary expert on working from home, it has the following very significant limitations when applied to the current situation.

1. The study was done on a Chinese travel company’s call center. What the effects of cultural differences between US and Chinese workers is hard to assess, but might be significant.
2. Call center employees are measured on fairly precise and straightforward metrics like number of calls handled per unit of time. Call center workers don’t really collaborate, innovate, or perform knowledge work, as it is typically defined.

3. The group chosen for the study volunteered to work from home. Subsequently the group was divided into a control group and a treatment group. However the initial choice of survey participants was biased by the selection method. It is interesting to note that at the conclusion of the study, approximately 50% of the WFH group chose to return to the office of their own volition.
4. The study was done in 2013.

These points don't invalidate the study on the subject it was designed to elucidate. It has these advantages:

1. It's unique in its methodology, which doesn't rely on questionnaires and self-assessment.
2. It was done by a well-regarded economist from Stanford.
3. It produced at least one simple-to-digest statistic, which was a 13% increase in productivity on the part of employees working from home.
4. However nothing about this study and its methodology really justifies its enormous utilization to prove things about post-COVID remote work.

The methodology point should cause readers and decision makers to consider whether inductive studies like this (more timely and more focused on US knowledge workers) should be given much more emphasis than survey-based data on workers' perceptions. There are a lot of studies that fit in this category, including: (Atkin, Chen and Popov, 2022), (Emanuel, Harrington and Pallais, 2023), (Gibbs, Mengel and Siemroth, 2021) and (Yang *et al.*, 2021).

Emanuel *et al* studied a software company which had a split workforce, thereby allowing the researchers to evaluate the broader question of proximity, the "parent" of the WFH discussion. They make the point that being in proximity actually reduced the amount of code produced by senior programmers (Emanuel, Harrington and Pallais, 2023, p. 4). On the other hand, proximity greatly improved the mentoring of younger and female coders, resulting in lower turnover. The authors also postulate that the code produced by the

proximate teams might be higher quality.

Yang *et al* worked with data captured from 61,000+ Microsoft employees, including emails, instant messages, and calendars, concluding that

“firm-wide remote work caused the collaboration network of workers to become more static and siloed, with fewer bridges between disparate parts. Furthermore, there was a decrease in synchronous communication and an increase in asynchronous communication. Together, these effects may make it harder for employees to acquire and share new information across the network.” (Yang et al., 2021, p. 1).

One of the most salient studies was a follow-on by Bloom himself (Bloom, Han and Liang, 2022) which combined self-assessment components with inductive data. There were many interesting conclusions of the study related to employee willingness to participate in the study², employee evaluations³, attrition rates, and lines of code generated⁴. It’s important to note this this study addressed hybrid work rather than fully remote workers. One of their conclusions was that there wasn’t a significant difference along several vectors between **hybrid** work and in-office work, with the exception of rates of attrition, which were lower for employees but higher for managers.

The utilization of this study is discussed below.

² *There were 1,612 employees selected for the study. After an initial email requested their participation, only 518 employees volunteered. The authors speculate this might be about perceptions that volunteers were less interested in working hard. After this low response, management required all 1,612 employees to participate. (Bloom, Han and Liang, 2022, p. 7)*

³ *Performance reviews and promotion rates were essentially the same between the groups. (Bloom, Han and Liang, 2022, p. 14)*

⁴ *A large percentage (70%) of the participants were coders and the hybrid workers produced about 4.4% more lines of code, which the authors characterize as statistically insignificant. (Bloom, Han and Liang, 2022, p. 14)*

3. Internal Corporate Research

The third category could be called “internal research,” and that is a type of research conducted inside organizations, not as part of an academic project, but to inform management in their decision-making. By its nature, this type of data is not published (or at least not published until well after it is gathered, analyzed, and employed by the organizations doing the research).

However we know two things about it, one with certainty and other through logical inference. It is certain that this type of data gathering is being done extensively inside organizations like Amazon, Google, Facebook and other similar firms (including probably financial service companies). The inference derives from looking at what these organizations are stating are their work from home policies, and the changes in those policies over time. These changes are front-page news in the business press, and at the risk of oversimplifying, it seems as though most organizations are trying to reduce the amount of working from home. Those “mandates” and the backlash associated with them are the core dialectic of most business press coverage, with headlines that often characterize the tension between return to office (RTO) policies and employee wishes as a “battle” or “war.”⁵



Andy Jassy, Amazon CEO

⁵ See for example (Goldberg, Emma, 2023) in the paper of record.

Self-assessment Research

I would argue that while self-assessment of data has some value, it has to be considered in terms of what is known about self-assessment generally, and there is a very broad and deep body of research on biases in self-assessment and motivated reasoning in general. Three articles that give insight into this (extremely broad) field are (Karpen, 2018), (Dunning *et al.*, 2003) and (Epley and Gilovich, 2016).

Karpen provides a broad overview of research, and makes the point that

“To date, few studies have found a strong or even moderate relationship between self assessment and actual ability.” (Karpen, 2018, p. 441)

He adds:

“One of the primary reasons that self-knowledge biases are so recalcitrant is that the mechanisms driving them operate below consciousness. It is difficult to modify processes of which we are unaware. While people have a multitude of cognitive tools for diffusing threatening information and enhancing positive information, the most common are self-serving reasoning, biased hypothesis testing, and biased recall.” (Karpen, 2018, p. 442)

The mechanism for this bias is typically called motivated reasoning, which is described in (Epley and Gilovich, 2016). They write:

“The crucial point is that the process of gathering and processing information can systematically depart from accepted rational standards because one goal–desire to persuade, agreement with a peer group, self-image, self-preservation–can commandeer attention and guide reasoning at the expense of accuracy.” (Epley and Gilovich, 2016, p. 135)

It isn't that the respondents in these studies are intentionally distorting their work results, but instead

“People’s motivations thus do not directly influence what they believe. Instead, their motivations guide what information they consider, resulting in favorable conclusions that seem mandated by the available evidence.” (Epley and Gilovich, 2016, p. 137)

Among the recent WFH studies, (Gavoille and Hazans, 2022) rely on self-assessment data for their work, but have the self-awareness to write

“...a common weakness of the literature focusing on the WFH/productivity is the reliance on self-reported productivity measures. Hence, one cannot fully rule out strategic manipulation or self deception when those who like (respectively, dislike) teleworking overstate (respectively, understate) their ‘from home’ productivity.” (Gavoille and Hazans, 2022, p. 3)

To be fair, they also observe that

“At the same time, there is no reason for respondents to misreport their preferences, so the estimates regarding willingness to keep working remotely do not suffer from self-reporting bias.” (Gavoille and Hazans, 2022, p. 3)

This comment points the way to use these studies: to assess *attitudes* about WFH, RTO, work-life balance, and mental health, rather than productivity.

Motivated Reasoning and Dunning-Kruger

As one might expect, the majority of motivated reasoning and Dunning-Kruger effect is found in the popular business press, including publications like *Business Insider*, *Fortune*, *Vox*, *The New York Times*, etc. Unfortunately much of the debate around work from home goes on in these publications.

One example of the Dunning effect is in a piece in *Business Insider* with the unfortunate title “Return to office mandates are ridiculous, unnecessary and, in some instances, even cruel.” The author writes:

“But there’s also research, like this Stanford study from 2014, that shows long hours in the office actually reduce productivity for most workers.” (Bort, 2023)

However, clicking on the link one finds that the research piece cited is in fact about munitions workers and the data was compiled by the British Health of Munion Workers Committee (HMWC) in 1915 and 1916 (Pencavel, 2014, p. 25). Does this hundred-year-old manufacturing labor data have any relevance to today? It is doubtful, and what is unarguable is that it doesn’t tell us much about remote knowledge work in 2023.

Bort’s main point stems from a comment made by Amazon CEO Andy Jassy, which she quotes from another *Business Insider* piece. The original piece described Jassy’s response declining giving up Amazon data about RTO and WFH, with the brush-off remark that it was a “judgement call” reinforced by Jassy’s discussions with managers and other CEOs. The quote from the original piece was:

“Amazon’s top leadership looked at” a number of pieces of data” over the past two years regarding remote work, Jassy said. Though he didn’t share any specific data, Jassy noted that the leadership team didn’t like the “actual results of our businesses” through a “significant chunk of time” during that period.”

“Other company CEOs helped Jassy make up his mind too. Jassy spoke to” 60 to 80 CEOs of other companies over the last 18 months,” and “virtually all of them” preferred in-office work, he said.” (Kim, 2023)

In Bort’s piece, this transforms into:

“What’s really going on has very little to do with productivity, where the data on the impact of office work versus remote is either mixed or non-existent. Even Amazon, a company that worships at the altar of data-driven decisions, has admitted that its RTO mandate isn’t based on data, but is a”-judgment” call.”(Bort, 2023)

As can be seen in the literature above, there is a **lot** of academic data on just that impact.

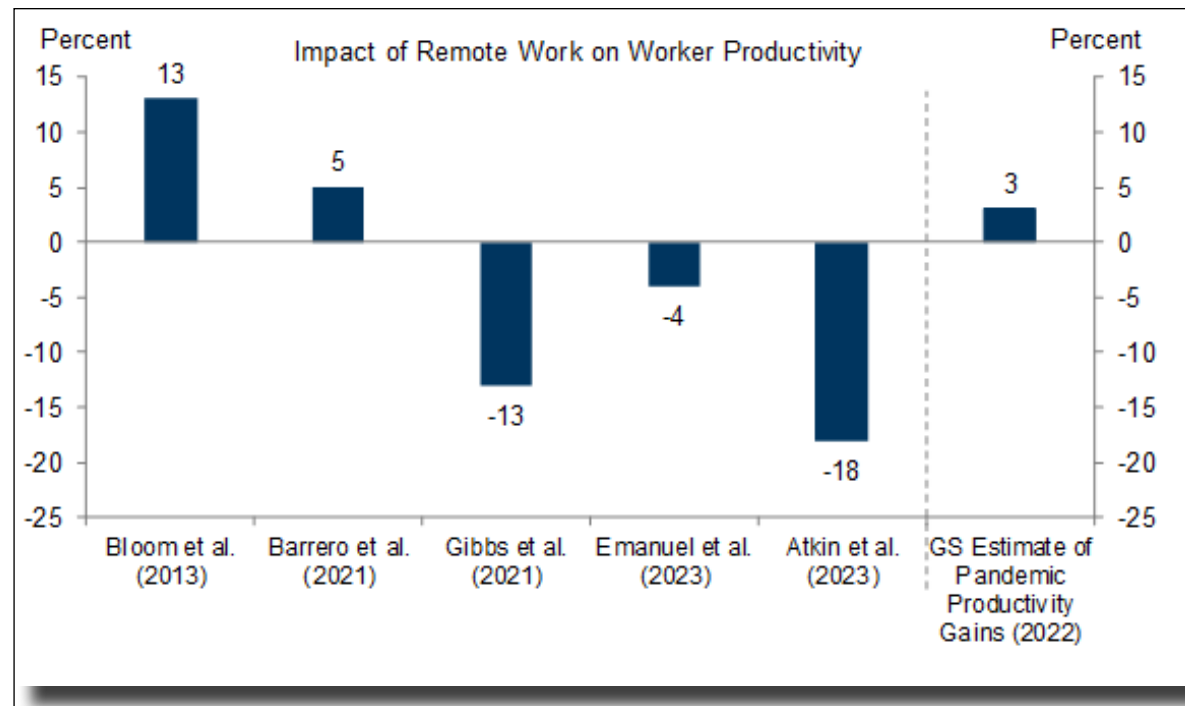
One of the key points about these anti-RTO pieces is that the authors had chosen, in some cases decades ago, to work from home. This work style works for them, and their unfavorable analysis of RTO mandates is a classic example of motivated reasoning, with a little Dunning-esque incompetence thrown in for good measure. To business writers who are committed to working from home..stand down on this one!

Not all examples of motivated reason in this sphere result from an inability to read and understand the research. Take for example a white paper prepared by the research group at Goldman Sachs (Goldman Sachs, 2023) entitled “Remote Work, Three Years Later,” initially cited in (Hoff, 2023) and (Goldman Sachs, 2023). Clearly Goldman’s research team is quantitatively sophisticated and aware of the academic studies. However, they write the following in their bullet point summary:

“Economic studies disagree on the productivity effects of WFH, with estimates ranging from -19% to +13% and compared to our baseline of +3% for pandemic productivity gains more generally.”

This phrase needs to be parsed carefully, and to assist this they provide the following chart:

Figure 1



The first column is from the original Bloom 2012 study, the timeliness and applicability to US office work having been questioned about above.

The second study in the chart, (Barrero, Bloom and Davis, 2021) states that:

“Using our survey data on self-assessed productivity effects of WFH, employer plans about who will work from home in the post-pandemic economy, commuting times and more, we estimate that the re-optimization over working arrangements in the post-pandemic economy will boost productivity by 4.6 percent relative to the pre-pandemic situation. The main source of this productivity boost is the savings in commuting time afforded by more WFH. This gain is missed by conventional productivity measures. Indeed, as conventionally measured, we project a productivity boost of only 1.0 percent in the post-pandemic economy.” (Barrero, Bloom and Davis, 2021, p. 4)

This paragraph is difficult to fully understand. The reason that “traditional productivity measures” miss the “gain” is that traditional productivity measures don’t count *commuting time* as *work*. Admittedly this is a complicated subject, because employee well-being might increase significantly if say 3 hours of weekly commuting is converted into 0 hours of commuting and 1.5 hours of work from home. However this doesn’t mean that they are more productive, only that some of their commuting has been replaced by work.⁶

Once this has been recomputed by the authors it becomes **1.0 percent**, which is surely inside of the standard error of the measurement.

Why did the Goldman researchers include a ten-year old Chinese data center study and why present the 2021 Barrero and Bloom data as a (rounded up) 5% productivity result?

Heading to the rightmost part of the chart, the motivation for the changes might be found. This final data element is labeled “GS Estimate of Pandemic Productivity Gains (2022).” In other words a prior year analysis by GS had come up with a positive change in productivity. What was the methodology for this analysis? Hard to tell unless you are a GS client, since the links provided above go to a GS client-only site.

If you removed the Bloom 2013 study, reported the Barrero 2021 result at 1%, and excluded the GS analysis, you would have four studies averaging -8.5%. Motivated reasoning?

⁶ *The basic measure of productivity employed by agencies like the Bureau of Labor Statistics is Productivity = Output / Hours. Thus if you produce 40 widgets in 40 hours of work your productivity is one widget / hour. If you include commuting in this calculation, two individuals with the same output, one commuting ten hours per week and other just five hours, the latter is therefore more productive than the former. This also implies that someone who converts commuting into work time is less productive if their output remains constant. As mentioned above, this is a complicated subject and therefore very subject to inappropriate simplification in the business press. Even academics seemed confused by this, for example (Stropoli, 2021). See the (BLS, 2023)*

Another good example of motivated reasoning can be found in (Tsipursky, 2023), also decrying the decisions of Amazon’s CEO.⁷ Echoing Kim and Bort above, he starts his piece with:

“Where are CEOs like Amazon’s Andy Jassy getting the data to inform their return-to-office policies? Unfortunately, too many are getting their data from the same place: an echo chamber of like-minded CEOs who use their feelings and intuitions to make these pivotal decisions.”

All these authors have latched onto Jassy’s remarks to draw the conclusion that Amazon—perhaps the most data-driven business imaginable—is making its decisions without data. As mentioned above, polling the experience of other CEOs, who presumably have a variety of sources of data in a variety of contexts, seems like a pretty good source of data, and is certainly not the product of “feelings” on the part of Jassy.

Tsipursky’s article uses the phrase “echo chamber” ten times in the short post. Un-ironically he counters the Jassy “echo chamber” with a description of an Amazon internal Slack channel which led to a petition in favor of remote work at Amazon:

“In the petition, Amazon employees argued, based on research, that remote work improved productivity, recruitment, work/life balance, inclusion efforts, and reduced corporate expenses.”

He doesn’t cite the research employed. And anyone who has ever used Slack should agree that in some ways it’s the ultimate echo chamber.

There’s a little Dunning-Kruger in this piece as well. He states:

“A whopping 80% of bosses reported that they regret their initial return-to-office decisions, according to new research from Envoy.”

⁷ To be fair to Fortune this piece is in their “commentary” section, albeit on their web site. At the bottom they have a disclaimer: *The opinions expressed in Fortune.com commentary pieces are solely the views of their authors and do not necessarily reflect the opinions and beliefs of Fortune.*

In fact that's not what the Envoy piece says at all. Their white paper instead says the following:⁸

“80% of executives say they would have approached their company’s return-to-office strategy differently if they had access to workplace data to inform their decision-making.”

The white paper cited isn't even really about remote work; it's predominantly about the value of Envoy's product for gathering workplace data:

“Envoy Workplace is the only fully integrated workplace solution for all your workplace needs. Empower your employees to find coworkers, rooms, desks, and deliveries, while you get an accurate and unified view of workplace occupancy and space usage data to make confident, cost-saving decisions.”

Tsipursky describes himself as “the office whisperer” in his bio on the *Fortune* commentary site, and he states:

*“At all the companies I worked with to help guide their RTO, whether it’s from the start of the process or to refine it after RTO was already launched, we started with a thorough survey of staff opinions about remote work and the return to office. That survey involves questions around their preferences on RTO, intent to stay with various versions of RTO, whether they would recommend working here to their peers given these versions, their productivity on individual and collaborative tasks at home vs. in the office, and similar questions on well-being, happiness, morale, stress, and so on. The survey is available in the appendix of my best-selling book, *Returning to the Office and Leading Hybrid and Remote Teams*.”*

This is a good combination of reliance on self-assessment and motivated reasoning on the part of the consultant-author.

⁸ Available here: (Envoy, 2023)

Concluding Thoughts

The purpose of this article was not to attack the work-from-home self-assessment literature. That research performs several important purposes, including assessing how much people prefer flexibility, how much they would be willing to trade in compensation for flexibility, and how willing they would be to leave their organization if a five-day RTO plan were to be implemented. (Gavoille and Hazans, 2022) is a good example of using self-reported data to report both attitudes about productivity and attitudes around remote work preferences.

These self-assessment studies also illustrate the dichotomy between worker and management perspectives on working from home. As has been observed in many articles, there is a significant difference between the attitudes of workers and management when it comes to remote work, with many articles belittling the managers and characterizing them as rigid, unresponsive to changing realities of work, and reductive in their thinking about productivity. All of these observations, however, seem to be *ad hominem* and not supported by any actual research.

On the other hand, waiting for a significant number of studies with control groups, un-biased samples of workers, and objective measures of productivity might leave us waiting for quite some time. What would be useful would be to have more research like (Yang *et al.*, 2021).

While inductive academic studies that shed more light on this subject will take time to arrive, it is important to observe that lots of data is being accumulated and analyzed by organizations like Amazon, Google, Apple—in fact all of the data-driven management teams are undoubtedly studying the relationship between in-office and remote workers' productivity. Whether they are willing to share this (other than implicitly through their actions and public statements) remains to be seen. But the actions of various executives (despite Bort, Kim, and Tsipursky's claims) make it clear that their interpretation of the preliminary data suggests, at the minimum, that 3/4 day-in-office hybrid work or RTO dominate pure (or predominant) WFH.

Nothing definitive can be determined currently, as noted in the opening paragraph. The final test will be the ruthless will of the marketplace, in which companies that work from the office predominantly versus those that have predominantly WFH policies (such as having employees only come to the office once a week or even once a month) win the battles of innovation, revenues and profits. There is no question that, at least in the current tight labor market, providing employee flexibility improves retention and perhaps the recruiting of employees grown accustomed to working several days a week at home. Whether that advantage persists as labor markets soften will only be seen when several years of financial results are reviewed, and so, as the title states, the debate continues.

Appendix A

As mentioned above, Angelici *et al* perform two types *objective* analysis in conjunction with describing and quantifying a great deal of self-assessment data. They compare “smart working,” defined as one day a week in which the employee gets to choose her location, versus traditional methods. The sample was of 310 workers.

One objective analysis constructs an OLS regression between the treatment group and the control to determine the number of days of leave taken: the treated group took 5.6 fewer days of leave over the nine-month period of the study. (p. 23). This is a material conclusion, although not earthshaking since it represents approximately a 3% improvement. It is also worth noting that this outcome is what would be expected since having the freedom to set one’s schedule (even to the limited extent of the study) should obviate the need for certain types of leave.

The authors also perform a less clear-cut analysis on the study’s objective productivity data. Their methodology is described as follows:

“We use objective measures of workers’ performance calculated monthly by the firm (e.g., the number of dossiers processed during the month)..” (p. 7)

and

“The firm provides...a monthly numeric indicator of objective productivity for each worker, built from the results of each worker in his/her own tasks.” (p. 17)

This numeric indicator is a bit of a mixed bag. This is not a fault of the authors because the data was gathered from information stored by the tested company, and the data is therefore heterogeneous. For example:

“...for 84.5% of them, the indicator corresponds to an absolute number, while for the other 15.5%, it is measured as a numeric change with respect to the measurement of the previous period.” (p. 17)

and:

“Both the absolute number and the change reflect the exact number of executed tasks (e.g., the number of procedures completed, the number of contracts concluded, the number of transactions performed, etc.), or the evaluation of the employee on a scale reported by the supervisor, or the compliance with deadlines (yes or no).” (p. 17)

They then go on to state:

“We consider each monthly variation; the variations can be positive, negative or null. We sum all variations and create a dichotomous variable that has the value of 1 if that sum is positive, thus indicating an overall improvement of productivity over the treatment period, and is zero otherwise. Subsequently, we perform a logistic regression for this dichotomous variable.” (p. 17)

They run the logistic regression and report out the odds-ratio, which provides the probability that an improvement in performance was correlated with the treatment versus the probability that the improvement was not related to the treatment. As presented in their Table 8 (p. 48), this ratio is between 2x and 2.2x (depending on whether any control variables such as age, gender, etc. are considered in the regression). The p-value of this result is $<.05$, indicating statistical significance. Seems conclusive that the smart-working treatment group is a lot more likely to be more productive.

However further consideration of their “compression” methodology brings up a significant concern. Since any cumulative improvement over the months of the study gets judged as a “one” in the regression, and any cumulative reduction gets judged a “zero,” it’s hard to assess the meaningfulness of the odds ratio.

A simplified numerical example makes this problem evident. To the right is some constructed data.

Assume this is from an experiment with 20 samples, half from the control group and half from the treatment group. Also assume that seven of the ten treatment results were improvements and three were not, and that five of the control group results were improvements and five were not.

Further assume to illustrate the point that the overall mean improvement in the treatment group was 1% and the mean decline in the control group was 25%. The average productivity of the group would be -9.4%. What would the odds ratio of this test be? The answer is 2.33, much like the ratio in Angelici *et al's* analysis.

The odds ratio test as they report it just states that the odds of some improvement in the treatment group was twice the odds of improvement in the control group, without giving insight into how much the treatment group changed nor how much the control group changed.

While this is a fanciful example, it shows that without the absolute improvement numbers it's hard to see how much net productivity was achieved with smart-working.

This critique might seem like nit-picking..but the authors repeatedly describe the approach as including *objective* data as well as an extensive set of questionnaire-based information provided by employees and management. They probably recognize some of the criticisms of self-assessment raised in this paper, and wish to supplement that with things not as subject to motivated reasoning on the part of the study participants.

Example Data: Productivity -> Logistic Regression

Input (Training Set=1)	Productivity	Outcome (Net positive = 1)
1	1%	1
1	1%	1
1	1%	1
1	1%	1
1	1%	1
1	1%	1
1	1%	1
1	1%	1
1	-25%	0
1	-25%	0
1	-25%	0
0	1%	1
0	1%	1
0	1%	1
0	1%	1
0	1%	1
0	-25%	0
0	-25%	0
0	-25%	0
0	-25%	0
0	-25%	0
Mean Productivity	-9.4%	

References

- Aksoy, C.G. *et al.* (2022) 'Working from Home Around the World', *Brookings Papers on Economic Activity*, pp. 1–66.
- Angelici, M. and Profeta, P. (2020) 'Smart-Working: Work Flexibility Without Constraints', *SSRN Electronic Journal* [Preprint]. Available at: <https://doi.org/10.2139/ssrn.3556304>.
- Atkin, D., Chen, K. and Popov, A. (2022) 'The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley', *NBER, WORKING PAPER* 30147, p. 52.
- Barrero, J.M., Bloom, N. and Davis, S.J. (2021) *Why Working From Home Will Stick*. w28731. National Bureau of Economic Research, p. 69. Available at: <https://doi.org/10.3386/w28731>.
- Bloom, N. *et al.* (2014) 'Does Working from Home Work? Evidence from a Chinese Experiment *', *The Quarterly Journal of Economics*, 130(1), pp. 165–218. Available at: <https://doi.org/10.1093/qje/qju032>.
- Bloom, N., Han, R. and Liang, J. (2022) *How Hybrid Working From Home Works Out*. w30292. Cambridge, MA: National Bureau of Economic Research. Available at: <https://doi.org/10.3386/w30292>.
- BLS (2023) *Concepts : Handbook of Methods: U.S. Bureau of Labor Statistics*. Available at: <https://www.bls.gov/opub/hom/msp/concepts.htm> (Accessed: 13 September 2023).
- Bort, J. (2023) *Return to office mandates are ridiculous, unnecessary and, in some instances, even cruel*. Business Insider. Available at: <https://www.businessinsider.com/return-to-work-mandates-rto-office-cruel-2023-9> (Accessed: 10 September 2023).
- Dunning, D. *et al.* (2003) 'Why People Fail to Recognize Their Own Incompetence', *Current Directions in Psychological Science*, 12(3), pp. 83–87. Available at: <https://doi.org/10.1111/1467-8721.01235>.

Emanuel, N., Harrington, E. and Pallais, A. (2023) *The Power of Proximity to Coworkers*. The Federal Reserve Bank of New York.

Envoy (2023) *Accurate data is core to the survival of the physical workplace*. Envoy. Available at: <https://envoy.com/content/datareport/>.

Epley, N. and Gilovich, T. (2016) 'The Mechanics of Motivated Reasoning', *Journal of Economic Perspectives*, 30(3), pp. 133–140. Available at: <https://doi.org/10.1257/jep.30.3.133>.

Gavoille, N. and Hazans, M. (2022) 'Personality Traits, Remote Work and Productivity', *IZA DP No. 15486* [Preprint]. Available at: <https://doi.org/https://dx.doi.org/10.2139/ssrn.4188297>.

Gibbs, M., Mengel, F. and Siemroth, C. (2021) *Work from Home & Productivity: Evidence from Personnel & Analytics Data on IT Professionals*. Rochester, NY. Available at: <https://doi.org/10.2139/ssrn.3843197>.

Goldberg, Emma (2023) *Meet the Return-to-Office Whisperers | Readwise*. nytimes.com. Available at: <https://www.nytimes.com/2023/04/11/magazine/return-to-office-consultants.html?smid=nyt-core-ios-share&referringSource=articleShare> (Accessed: 16 September 2023).

Goldman Sachs (2023) *Remote Work, Three Years Later*. Available at: <https://www.gspublishing.com/content/research/en/reports/2023/08/28/6fd0a8a0-3831-4ace-b577-a93337f01ec8.html> (Accessed: 11 September 2023).

Herbsleb, J.D. and Mockus, A. (2003) 'An empirical study of speed and communication in globally distributed software development', *IEEE Transactions on Software Engineering*, 29(6), pp. 481–494. Available at: <https://doi.org/10.1109/TSE.2003.1205177>.

Hoff, M. (2023) *No one can agree on whether remote work is making employees more productive*. Business Insider. Available at: <https://www.businessinsider.com/remote-work-wfh-productivity-impact-no-consensus-goldman-sachs-2023-8> (Accessed: 16 September 2023).

- Karpen, S.C. (2018) 'The Social Psychology of Biased Self-Assessment', *American Journal of Pharmaceutical Education*, 82(5), pp. 441–448.
- Kim, E. (2023) *In leaked recording, Amazon CEO Andy Jassy wouldn't share data for his controversial RTO mandate. Said it was a 'judgement' call, like launching AWS*. Business Insider. Available at: <https://www.businessinsider.com/amazon-andy-jassy-no-data-return-to-office-rto-aws-2023-8> (Accessed: 10 September 2023).
- Kitagawa, R. *et al.* (2021) 'Working from home and productivity under the COVID-19 pandemic: Using survey data of four manufacturing firms', *PLOS ONE*. Edited by S.A. Useche, 16(12), p. e0261761. Available at: <https://doi.org/10.1371/journal.pone.0261761>.
- Morikawa, M. (2022) 'Work from home productivity during the COVID 19 pandemic: Evidence from Japan', *Economic Inquiry*, 60(2), pp. 508–527. Available at: <https://doi.org/10.1111/ecin.13056>.
- Pencavel, J. (2014) 'The Productivity of Working Hours', *IZA Discussion Paper No. 8129* [Preprint].
- Smite, D. *et al.* (2022) 'Changes in perceived productivity of software engineers during COVID-19 pandemic: The voice of evidence', *Journal of Systems and Software*, 186, p. 111197. Available at: <https://doi.org/10.1016/j.jss.2021.111197>.
- Stropoli, R. (2021) 'Are we really more productive working from home?', *Chicago Booth Rev* [Preprint].
- Tsipursky, G. (2023) *Amazon's Andy Jassy shouldn't make return-to-office decisions in an echo chamber of CEOs driven by feelings*. Fortune. Available at: <https://fortune.com/2023/09/05/amazon-andy-jassy-return-to-office-decisions-echo-chamber-ceo-feelings-work-gleb-tsipursky/> (Accessed: 13 September 2023).
- Yang, L. *et al.* (2021) 'The effects of remote work on collaboration among information workers', *Nature Human Behaviour*, pp. 1–12. Available at: <https://doi.org/10.1038/s41562-021-01196-4>.