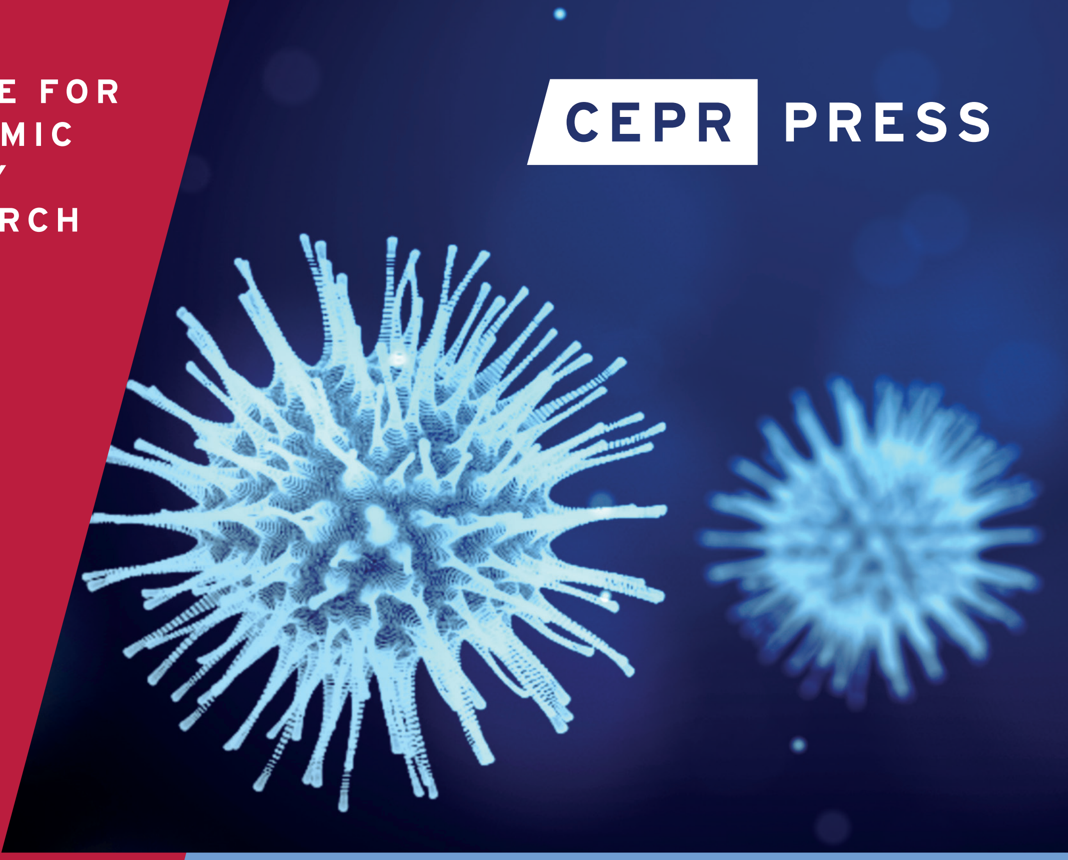
CENTRE FOR
ECONOMIC
POLICY
RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 82
23 JUNE 2021

Covid Economics
Vetted and Real-Time Papers

Covid Economics, Vetted and Real-Time Papers, from CEPR, brings together formal investigations on the economic issues emanating from the Covid outbreak, based on explicit theory and/or empirical evidence, to improve the knowledge base.

Founder: Beatrice Weder di Mauro, President of CEPR
Editor: Charles Wyplosz, Graduate Institute Geneva and CEPR

Contact: Submissions should be made at https://portal.cepr.org/call-papers-covid-economics by 23 June 2021.

Other queries should be sent to covidecon@cepr.org.

The Centre for Economic Policy Research (CEPR)

The Centre for Economic Policy Research (CEPR) is a network of over 1,500 research economists based mostly in European universities. The Centre's goal is twofold: to promote world-class research, and to get the policy-relevant results into the hands of key decision-makers. CEPR's guiding principle is 'Research excellence with policy relevance'. A registered charity since it was founded in 1983, CEPR is independent of all public and private interest groups. It takes no institutional stand on economic policy matters and its core funding comes from its Institutional Members and sales of publications. Because it draws on such a large network of researchers, its output reflects a broad spectrum of individual viewpoints as well as perspectives drawn from civil society. CEPR research may include views on policy, but the Trustees of the Centre do not give prior review to its publications. The opinions expressed in this report are those of the authors and not those of CEPR.

Editorial Board

Ethics

Covid Economics will feature high quality analyses of economic aspects of the health crisis. However, the pandemic also raises a number of complex ethical issues. Economists tend to think about trade-offs, in this case lives vs. costs, patient selection at a time of scarcity, and more. In the spirit of academic freedom, neither the Editors of *Covid Economics* nor CEPR take a stand on these issues and therefore do not bear any responsibility for views expressed in the articles.

Submission to professional journals

The following journals have indicated that they will accept submissions of papers featured in *Covid Economics* because they are working papers. Most expect revised versions. This list will be updated regularly.

American Economic Journal, Applied Economics
American Economic Journal, Economic Policy
American Economic Journal, Macroeconomics
American Economic Journal, Microeconomics
American Economic Review
American Economic Review, Insights
American Journal of Health Economics
Canadian Journal of Economics
*Econometrica**
Economic Journal
Economics Letters
Economics of Disasters and Climate Change
International Economic Review
Journal of Development Economics
*Journal of Econometrics**
Journal of Economic Growth
Journal of Economic Theory
*Journal of the European Economic Association**
Journal of Finance
Journal of Financial Economics
Journal of Health Economics
Journal of International Economics
*Journal of Labor Economics**
Journal of Monetary Economics
Journal of Public Economics
Journal of Public Finance and Public Choice
Journal of Political Economy
Journal of Population Economics
Quarterly Journal of Economics
*Review of Corporate Finance Studies**
Review of Economics and Statistics
*Review of Economic Studies**
Review of Financial Studies

(*) Must be a significantly revised and extended version of the paper featured in *Covid Economics*.

Covid Economics
Vetted and Real-Time Papers

Issue 82, 23 June 2021

Contents

Covid Economics **Issue 82, 23 June 2021**

Restaurant closures during the pandemic: A descriptive analysis[1]

Dmitry Sedov[2]

Date submitted: 11 June 2021; Date accepted: 19 June 2021

In this paper, I describe the restaurant business closure patterns in the year 2020. I use Yelp data collected first in late 2019 and then in early 2021 to study restaurant and location characteristics related to the exit decisions. I find that higher rated restaurants as well as restaurants located further away from central city areas were less likely to close during 2020.

Covid Economics 82, 23 June 2021: 1-9

1 I thank Timur Abbiasov, Anna Algina, Gaston Illanes, Robert Porter and Mar Reguant for helpful comments and discussions. Special thanks to Jonathan Wolf and the SafeGraph team for access to data, clarifications and thoughtful remarks.
2 PhD Student, Department of Economics, Northwestern University.

1 Introduction

During 2020, the first year of the COVID-19 pandemic in the US, restaurants suffered from reduced consumer traffic due to multiple factors: lockdowns, operations restrictions and social distancing. Which restaurants were more likely to exit the industry in this challenging time for the industry? In the present paper I provide descriptive evidence on this question in the context of major US urban areas using data from the review platform Yelp and the location data company SafeGraph. Specifically, I explore location- and restaurant-specific characteristics that explain variation in restaurant closure decisions.

The data on urban restaurants used in this paper was gathered from Yelp in two rounds: in late 2019 and early 2021 and is thus well suited for research on business closures in 2020. The dataset includes information on key restaurant characteristics, as well as each restaurant's open/closed status at the end of the observation period, which enables shedding light on the factors related to restaurant exit decisions. Using this dataset, I document the across-cities differences in restaurant exit rates, which range from 9.6% in El Paso to 21.5% in Honolulu. Next, leveraging Yelp and SafeGraph data, I build descriptive statistics on the restaurant closure patterns. Specifically, I estimate the binary response econometric models and report the Average Partial Differences that summarize the association between restaurant characteristics and exit. I find that higher rating score and review count are robustly associated with lower restaurant exit probabilities. A 1-star increase in the restaurant's rating is associated with a roughly 1.2% lower chance of restaurant closure. Additional 100 reviews at the beginning of the observation period are associated with $0.9-1.8\%$ lower probability of restaurant exit. Also, restaurants relying on the foot traffic generated due to their within-city location were relatively less likely to survive the pandemic year.

By studying restaurant closure patterns during the COVID-19 pandemic I contribute to several strands of literature. First, I expand a relatively small body of research specifically aimed at studying restaurant exit decisions. Luo and Stark (2014) establish that the median restaurant lifespan is 4.5 years and find that 17% of restaurants fail during the first year of existence. An earlier study by Parsa et al. (2005) found a 26% failure rate in the first year and a decreasing failure rate pattern over time. Follow-up studies include Parsa et al. (2011), Parsa et al. (2011), Parsa et al. (2015), Parsa et al. (2019). Closely related is also the work by Tao and Zhou (2020) who use similar data as I do in this paper, but concentrate on the prediction task rather than a descriptive one. Second, the present paper is also related to the industry dynamics literature in IO that, among other issues deals with firm entry and exit decision as well as the turnover on the market level. Geroski (1995), Sutton (1997) and Caves (1998) provide excellent surveys of the early empirical findings based on across-industry studies, including evidence on the variation of entry/exit rates and the relationship between market features and turnover. The early evidence on firm survival has been supported (see, e.g. Audretsch et al. (2000) and Fackler et al. (2013)) and also expanded (Agarwal and Audretsch (2001), Agarwal and Gort (2002), Klepper (2002), Bernard and Jensen (2007)) by subsequent papers. Firm exit decisions have also been studied using structural models of firm behavior. Examples include Aguirregabiria and Mira (2007), Ryan (2012), Dunne et al. (2013), Yang (2013) and Fowlie et al. (2016). While this latter strand of literature allows for studying policy counterfactuals, the goal of this paper is more in line with the former strand of descriptive evidence: understanding the factors related to exit decisions in the restaurant industry during the COVID-19 pandemic. To this end, this paper is also related to the small but growing research on business disruptions during the pandemic (e.g. Bartik et al. (2020), Fairlie (2020), Koren and Peto (2020) or Crane et al. (2021)).

Covid Economics 82, 23 June 2021: 1-9

	% NA	*Q10*	*Q25*	*Med*	*Q75*	*Q90*	*Mean*	*SD*
Closed	0.00	0.00	0.00	0.00	0.00	1.00	0.15	0.36
Price	16.61	1.00	1.00	2.00	2.00	2.00	1.57	0.61
Rating	0.00	2.50	3.00	4.00	4.00	4.50	3.59	0.88
Reviews	0.00	4.00	14.00	54.00	177.00	425.00	168.50	365.25
# categories	0.00	1.00	1.00	2.00	3.00	3.00	1.99	0.84
City center dist. (km)	0.00	1.27	3.23	7.30	12.89	18.66	8.89	7.44
Est. nearby	0.00	17.00	33.00	63.00	137.00	332.00	136.84	210.48

Table 1: Restaurants dataset summary statistics. Number of observations: 128,285.

The rest of the paper is structured as follows. Section 2 presents key facts about the data. Section 3 describes the econometric analysis of factors related to restaurant exit decisions. Section 4 concludes.

Covid Economics 82, 23 June 2021: 1-9

2 Data

Two data sources are used for the analysis discussed in this paper. First, I use the data collected from the Yelp restaurant review platform. This data provides information on restaurant characteristics and exit decisions. Second, I use data from the location data company SafeGraph to construct additional covariates related to restaurant location characteristics. The data obtained by combining information from these two sources covers 128,285 restaurants in 42 major US cities.

The timing of data collection allows me to concentrate on the [first] year of the COVID-19 pandemic. The data on restaurants' names, locations and characteristics were first collected in late 2019 using a scraping routine that systematically parsed Yelp Fusion API[1]. The second round of data collection was done in early 2021, using the previously gathered set of unique Yelp restaurant identifiers. The target element of interest during the second round was the *restaurant-closed* indicator, which I view as the ground truth on restaurant exit for the purposes of this paper[2].

The second source of data is the location data company SafeGraph, which collects information on points-of-interest in the US (defined as places outside of home where people spend time and/or money). For the purposes of this paper, it is most important to note that the SafeGraph data allows me to observe the locations of roughly 4.4 mln establishments across multiple industries and to quantify the proximity of restaurants to other businesses (see Abbiasov and Sedov (2021) or Sedov (2021) for a more detailed description of the dataset). Specifically, using the extract of SafeGraph's data from July 2019, I use the counts of nearby establishments in the 500 meter radies of each sample restaurant to quantify the likely restaurant reliance on the traffic generated by nearby establishments. Table 1 provides the summary statistics on the resulting dataset.

Several facts about the data are worth stating. 15.2% of restaurants in the sample closed business during the year 2020. To illustrate the geographic variation in restaurant closures, Figure 1 depicts the exit rates across sample cities. Honolulu featured the highest exit rate of 21.5% among the sample

[1]The initial set of restaurants was obtained from the SafeGraph database of points-of-interest and complemented with a search for "food" around a dense grid of points corresponding to Census Block Group centroids.

[2]Yelp describes this field as indicating "whether business has been (permanently) closed" but is not explicit whether the field is self-reported or crowd-sourced from the platform users.

Covid Economics 82, 23 June 2021: 1-9

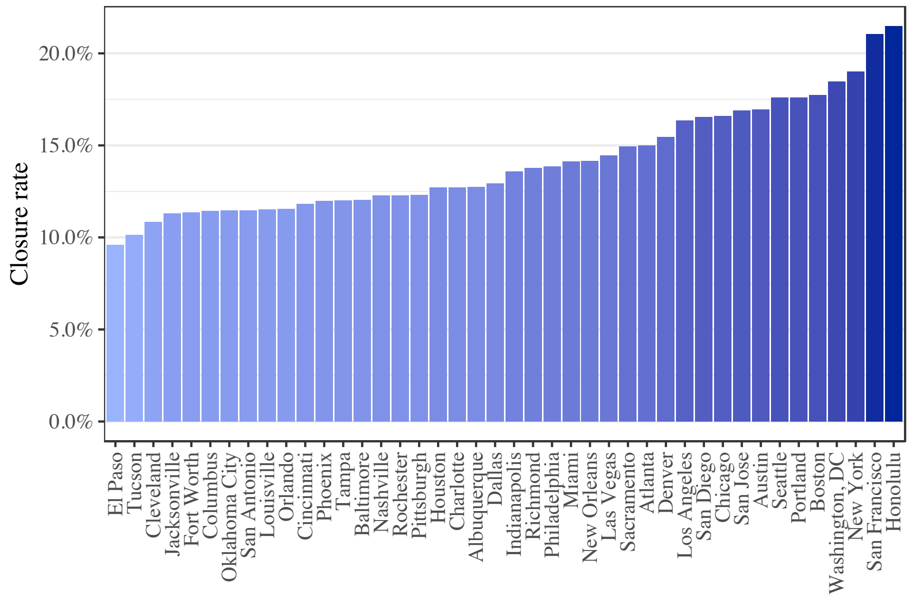


Figure 1: Restaurant closure rates across major US cities between late 2019 and early 2021.

cities, while El Paso's exit rate was the lowest at 9.6%. Figure 2 displays the relationship between market size (measured as restaurant count on the city level) and restaurant closure rates. Larger markets have experienced higher restaurant exit rates: a 1000-increase in restaurant count is associated with a 0.46% increase in the overall restaurant closure rate.

3 Exit-related factors

To understand the role of different factors shaping restaurant exit decisions, I estimate binary response models (LPM, logit and probit) linking restaurant characteristics and closure. In my empirical specification, restaurant characteristics include variables that are related to both the features of a restaurant itself and to the features of its location. Restaurant-specific characteristics include dummies for price categories, Yelp rating score and review count, primary cuisine category and the total number of cuisine categories associated with the restaurant. Restaurant location features consist of city dummies, latitude and longitude, the count of nearby establishments as well as the distance from city center.

Table 2 presents the coefficient estimates for the alternative binary response variables. Column (1) represents the LPM with city and cuisine category fixed effects, while column (2) represents the LPM with the interacted city-cuisine fixed effects. Columns (3)-(4) represent logit models with the same alternative sets of fixed effects. Columns (5)-(6) show the probit estimates. The signs of the coefficient estimates are the same across specifications for all variables of interest.

I first discuss the restaurant characteristics coefficient estimates. The coefficient on the $-dummy is negative, indicating that, relative to the missing price label, $-priced restaurants were less likely to

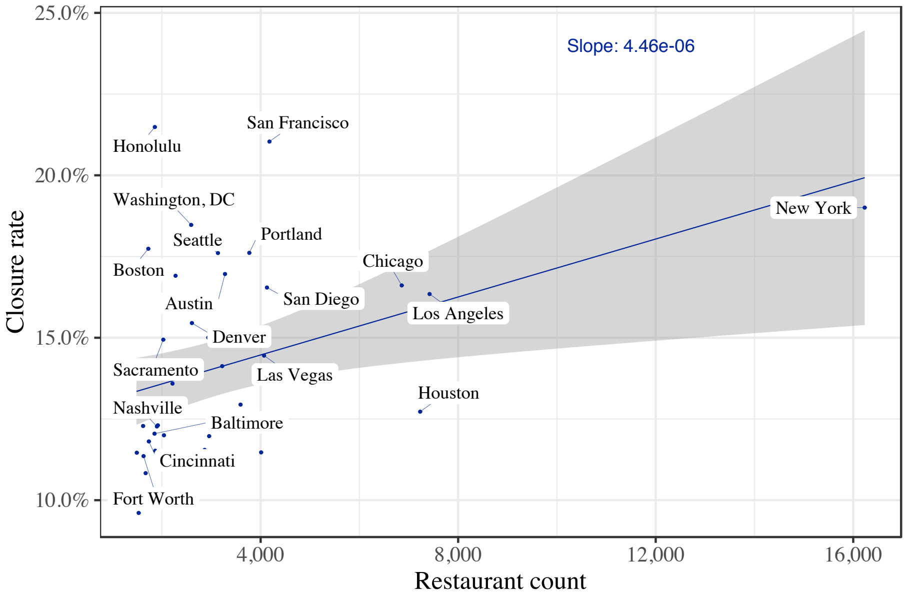


Figure 2: Restaurant closure rates by market size (restaurant count).

close in 2020. The coefficients on $$, $$$ and $$$$ are positive: higher-priced restaurants were more likely to close relative to the baseline. The coefficient on the $$$$-dummy, however, is not significant at the 5% level across all specifications. The rating and review count coefficients are negative and significant, implying that higher-quality and more frequently reviewed restaurants were less likely to close during the observation period. The total number of cuisine categories was estimated to have negative coefficients: restaurants with more diverse food were more likely to survive during 2020.

Several coefficients on the location-specific restaurant features provide additional insight. The coefficient on the count of nearby establishments is positive, indicating that restaurants that are located close to many other businesses were more likely to close. These restaurants likely relied on the foot traffic generated by the nearby establishments, and probably suffered relatively more from the pandemic, which is one of the channels that could result in higher exit rates among such restaurants. The negative coefficient on the variable measuring the distance from the city center indicates that centrally located restaurants were more likely to exit the business. Again, this may be related to a relatively higher fall of foot traffic to central city areas during the 2020 pandemic.

Additionally, it is worth noting that the alternative sets of fixed effects do not appear to dramatically change any of the coefficient estimates. A likely explanation is that exit rates across cities and within a cuisine category are not very different. Since the comparison of the coefficient estimates between alternative models is not very informative, I report the Average Partial Differences (APDs) corresponding to the variables of interest in Table 3. APDs describe the average change in the probability of restaurant closure conditional on a marginal increase in the respective variable (or a change from the baseline to the target value in case of a categorical variable). Formally, an Average Partial Difference is defined exactly as an Average Partial Effect (see Wooldridge (2010)),

CEPR PRESS

	Dependent variable: restaurant closed					
	LPM		*Logit*		*Probit*	
	(1)	(2)	(3)	(4)	(5)	(6)
\$	−0.020	−0.020***	−0.118***	−0.115***	−0.077***	−0.076***
	(0.017)	(0.005)	(0.025)	(0.025)	(0.013)	(0.014)
\$\$	0.012	0.012*	0.207***	0.206***	0.096***	0.095***
	(0.011)	(0.006)	(0.024)	(0.025)	(0.014)	(0.014)
\$\$\$	0.016	0.016	0.269***	0.273***	0.127***	0.128***
	(0.012)	(0.009)	(0.047)	(0.048)	(0.027)	(0.027)
\$\$\$\$	0.002	0.003	0.145	0.161	0.061	0.067
	(0.014)	(0.014)	(0.096)	(0.097)	(0.054)	(0.055)
Rating	−0.014**	−0.014***	−0.099***	−0.103***	−0.054***	−0.056***
	(0.005)	(0.002)	(0.010)	(0.010)	(0.006)	(0.006)
Reviews (100s)	−0.009***	−0.009***	−0.140***	−0.143***	−0.067***	−0.068***
	(0.001)	(0.000)	(0.005)	(0.005)	(0.002)	(0.002)
Est. nearby (100s)	0.012***	0.012***	0.069***	0.069***	0.042***	0.042***
	(0.003)	(0.002)	(0.004)	(0.004)	(0.002)	(0.003)
City center dist. (km)	−0.002***	−0.002***	−0.024***	−0.024***	−0.012***	−0.012***
	(0.000)	(0.000)	(0.002)	(0.002)	(0.001)	(0.001)
# categories	−0.007	−0.007***	−0.034***	−0.035***	−0.020***	−0.021***
	(0.005)	(0.001)	(0.010)	(0.010)	(0.005)	(0.006)
City FE	✓		✓		✓	
Category FE	✓		✓		✓	
City-Category FE		✓		✓		✓
Observations	128,281	128,281	128,281	128,281	128,281	128,281
R^2	0.026	0.033				
Adjusted R^2	0.025	0.026				
Log Likelihood			−52765.3	−52217.4	−52811.5	−52267.0
Note:						*p<0.05; **p<0.01; ***p<0.001

Table 2: Coefficient estimates for the binary response models. Standard errors clustered at the FE levels for the LPM models. Latitude and longitude were included as covariates in all of the specifications but were omitted from the table; the corresponding coefficient estimates were insignificant. 4 observations were omitted from the analysis due to missing latitude / longitude.

but substituting *effect* with *difference* since the the estimates of this paper are meant to be purely descriptive.

Table 3 allows comparing the estimation results across the models. The estimated APDs appear to be of similar magnitude in the LPM, logit and probit models. The change from the reference price category (missing) to the \$-category is associated with a 1.4-2% drop in the closure probability. In turn, the change from the baseline to \$\$ category is associated with a 1.2-2.3% increase in the probability of restaurant exit. A restaurant with an extra star of the rating score can be expected to have a 1.2-1.4% higher probability of survival. A restaurant with an additional cuisine category reported is observed to be 0.4-0.7% less likely to close. The review count (in 100s) and the number of nearby establishments (also in 100s) exhibit a relatively weaker association with the restaurant survival probability with the respective APDs being around −0.1 and 1.8% depending on the specification. Finally, a 1-km increase

Covid Economics 82, 23 June 2021: 1-9

	Response: probability of restaurant closing					
	LPM		*Logistic*		*Probit*	
	(1)	(2)	(3)	(4)	(5)	(6)
\$	−0.020*	−0.020***	−0.014***	−0.014***	−0.017***	−0.017***
	(0.017)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
\$\$	0.012*	0.012*	0.027***	0.026***	0.023***	0.022***
	(0.011)	(0.006)	(0.003)	(0.003)	(0.003)	(0.003)
\$\$\$	0.016*	0.016*	0.036***	0.037***	0.030***	0.031***
	(0.012)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)
\$\$\$\$	0.002	0.003	0.018*	0.021*	0.014*	0.016*
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)
Rating	−0.014**	−0.014***	−0.012***	−0.013***	−0.012***	−0.013***
	(0.005)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Reviews (100s)	−0.009***	−0.009***	−0.018***	−0.018***	−0.015***	−0.015***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Est. nearby (100s)	0.012***	0.012***	0.009***	0.009***	0.010***	0.010***
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
City center dist. (km)	−0.002***	−0.002***	−0.003***	−0.003***	−0.003***	−0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# categories	−0.007*	−0.007***	−0.004***	−0.004***	−0.005***	−0.005***
	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
City FE	✓		✓		✓	
Category FE	✓		✓		✓	
City-Category FE		✓		✓		✓

Table 3: Average Partial Differences for the binary response models. Standard errors clustered at the FE level in LPM specifications.

in the distance from the city center is associated with roughly 0.2-0.3% lower closure probability. Overall, the estimates produced by the standard econometric models are intuitive in both size and magnitude.

4 Conclusion

This paper describes the exit patterns of businesses in the notoriously high-turnover restaurant industry during 2020, the year of the unusual negative shock to the industry. Using data from Yelp and SafeGraph, I determine the restaurant- and location- specific factors related to closure decisions.

References

Abbiasov, T. and D. Sedov (2021). Do local businesses benefit from stadiums? The case of major professional sports leagues arenas. Technical report, Northwestern University.

Agarwal, R. and D. B. Audretsch (2001). Does entry size matter? the impact of the life cycle and technology on firm survival. *The Journal of Industrial Economics 49*(1), 21–43.

Agarwal, R. and M. Gort (2002). Firm and product life cycles and firm survival. *American Economic Review 92*(2), 184–190.

Aguirregabiria, V. and P. Mira (2007). Sequential estimation of dynamic discrete games. *Econometrica 75*(1), 1–53.

Audretsch, D. B., P. Houweling, and A. R. Thurik (2000). Firm survival in the Netherlands. *Review of Industrial Organization 16*(1), 1–11.

Bartik, A. W., M. Bertrand, Z. Cullen, E. L. Glaeser, M. Luca, and C. Stanton (2020). The impact of covid-19 on small business outcomes and expectations. *Proceedings of the National Academy of Sciences 117*(30), 17656–17666.

Bernard, A. B. and J. B. Jensen (2007). Firm structure, multinationals, and manufacturing plant deaths. *The Review of Economics and Statistics 89*(2), 193–204.

Caves, R. E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature 36*(4), 1947–1982.

Crane, L. D., R. A. Decker, A. Flaaen, A. Hamins-Puertolas, and C. Kurz (2021). Business exit during the covid-19 pandemic: Non-traditional measures in historical context.

Dunne, T., S. D. Klimek, M. J. Roberts, and D. Y. Xu (2013). Entry, exit, and the determinants of market structure. *The RAND Journal of Economics 44*(3), 462–487.

Fackler, D., C. Schnabel, and J. Wagner (2013). Establishment exits in Germany: the role of size and age. *Small Business Economics 41*(3), 683–700.

Fairlie, R. W. (2020). The impact of covid-19 on small business owners: Continued losses and the partial rebound in may 2020. Technical report, National Bureau of Economic Research.

Fowlie, M., M. Reguant, and S. P. Ryan (2016). Market-based emissions regulation and industry dynamics. *Journal of Political Economy 124*(1), 249–302.

Geroski, P. A. (1995). What do we know about entry? *International Journal of Industrial Organization 13*(4), 421–440.

Klepper, S. (2002). Firm survival and the evolution of oligopoly. *RAND journal of Economics*, 37–61.

Koren, M. and R. Peto (2020, April). Business disruptions from social distancing. *Covid Economics 1*(2), 13–31.

Luo, T. and P. B. Stark (2014). Only the bad die young: Restaurant mortality in the western US.

Parsa, H., A. Gregory, and M. Terry (2011). Why do restaurants fail? Part III: An analysis of macro and micro factors.

Parsa, H., J. C. Kreeger, J.-P. van der Rest, L. Xie, and J. Lamb (2019). Why restaurants fail? Part V: Role of economic factors, risk, density, location, cuisine, health code violations and GIS factors. *International Journal of Hospitality & Tourism Administration*, 1–26.

Parsa, H., J. Self, S. Sydnor-Busso, and H. J. Yoon (2011). Why restaurants fail? Part II: The impact of affiliation, location, and size on restaurant failures: Results from a survival analysis. *Journal of Foodservice Business Research 14*(4), 360–379.

Parsa, H., J.-P. I. van der Rest, S. R. Smith, R. A. Parsa, and M. Bujisic (2015). Why restaurants fail? Part IV: The relationship between restaurant failures and demographic factors. *Cornell Hospitality Quarterly 56*(1), 80–90.

Parsa, H. G., J. T. Self, D. Njite, and T. King (2005). Why restaurants fail. *Cornell Hotel and Restaurant Administration Quarterly 46*(3), 304–322.

Ryan, S. P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica 80*(3), 1019–1061.

Covid Economics 82, 23 June 2021: 1-9

Sedov, D. (2021). How efficient are firm location configurations? Empirical evidence from the food service industry. Technical report, Northwestern University.

Sutton, J. (1997). Gibrat's legacy. *Journal of Economic Literature 35*(1), 40–59.

Tao, J. and L. Zhou (2020). Can online consumer reviews signal restaurant closure: A deep learning-based time-series analysis. *IEEE Transactions on Engineering Management*, 1–15.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Yang, N. (2013). March of the chains: Herding in restaurant locations. *NET Institute*, 11–16.

Covid Economics 82, 23 June 2021: 1-9

Underreporting child maltreatment during the pandemic: Evidence from Colorado[1]

Alexa Prettyman[2]

Date submitted: 10 June 2021; Date accepted: 17 June 2021

As a result of the COVID-19 pandemic schools closed abruptly in March 2020 and Colorado issued a stay-at-home order during the month of April. Subsequently, child maltreatment reporting dropped by 31 percent. This paper documents the decline in referrals and reports during the year in Colorado and predicts counterfactual estimates using two strategies. One strategy assumes the underlying behavior for child maltreatment was unchanged from 2019 to 2020. This approach implies that about 30,000 referrals went unreported over the year as a result of the pandemic. The second strategy assumes the economic distress brought about by the pandemic altered the underlying prevalence of child maltreatment. In this case, over 38,800 cases of child maltreatment might have gone undetected. Scaling these results to the national level suggests millions of child maltreatment referrals went unreported. I find that the largest reduction in reporting comes from the stay-at-home order, followed by school closings. Moreover, there is suggestive evidence that these missed children are suffering from neglect and not abuse. These findings quantify another hardship brought about by the pandemic, underreporting, and underscore the role mandatory reporters play in detecting child maltreatment.

Covid Economics 82, 23 June 2021: 10-48

1 Alexa Prettyman wrote this draft during her doctoral studies at Georgia State University. Starting July 2021, Alexa's new role is senior statistician and research supervisor at the California Center for Population Research at UCLA. Refer to her website for the latest contact information.

2 Department of Economics, Andrew Young School of Policy Studies, Georgia State University.

1. Introduction

At the beginning of the COVID-19 pandemic, headlines across the United States read "Child abuse hotline calls are down during COVID-19, but abuse fears are up" and "More than 60% drop in calls to child abuse hotline spark safety concerns" (Callahan & Mink, 2020; Quander, 2020). State agencies across the country were reporting that child abuse and neglect reports dropped drastically, but they cautioned that the decline was not necessarily a function of reduced maltreatment, and instead a function of reduced reporting.[1] Another headline read, "Advocates express concerns about children falling through the cracks" (WCTV, 2020). In Colorado, I find that reporting decreased by 15 percent in 2020 relative to 2019. The biggest drop in reporting, of 31 percent, occurred between April and June, but reporting remained 14 to 18 percent below 2019 levels for the remainder of the year.

Over the past year, child maltreatment research has shown that overall fewer allegations of maltreatment were reported than expected in March and April (Baron et al., 2020; Rapoport et al., 2020; Weiner et al., 2020), school closures drastically reduced the number of cases detected (Baron et al., 2020, Cabrera-Hernández & Padilla-Romo, 2020), and stay-at-home orders increased the incidence of neglect (Bullinger et al., 2020). All of these studies use different methods and data, yet come to the same conclusion in line with the concerns expressed by news articles: potential victims of child abuse and neglect are going unnoticed. Baron and colleagues (2020) use real-time data from Florida to estimate that a total of 212,500 allegations across the US, 40,000 of which would have been substantiated, went unreported in March and April of 2020 as a result of school closures. In this paper, I use real-time data from Colorado to provide an updated national estimate on the number of unreported allegations and victims for the entire year. I find that millions of allegations may have gone unreported, potentially impacting over 100,000 victims during the year. In addition, I estimate how child maltreatment incidences and reporting changed as a result of the COVID-19 pandemic, school closures, and the stay-at-home order. Not surprising, all three events concurrently resulted in the largest decline in reporting, followed by the pandemic-induced school closures.

[1] This concern is not unique to the United States. Headlines in Canada read "Child protection reports on P.E.I. climb despite fewer eyes amid COVID."

There are three main contributions of this paper. First, this paper uses real-time child maltreatment data and a model, similar to Baron et al. (2020), to predict the number of calls that would have been made to the Child Protective Services' (CPS) hotline in 2020 had the pandemic not occurred. The counterfactual number of child maltreatment referrals is calculated two ways. The simplest way is by assuming that referrals would have followed a similar pattern in 2020 as previous years. Alternatively, the pandemic limited interactions between children and mandatory reporters through school closings and stay-at-home orders, and increased child maltreatment risk factors, such as unemployment, parental burnout, and adverse coping mechanisms, like alcohol abuse.[2] For this reason, a second counterfactual is estimated taking into account the rise in unemployment and alcohol consumption. Comparing the two counterfactuals and the observed number of referrals sheds insight onto the two mechanisms in which the pandemic impacted child maltreatment. The first counterfactual underscores the importance of mandatory reporters. The second counterfactual demonstrates how economic hardships and coping mechanisms brought about by the pandemic contribute to child maltreatment.

As a second contribution, this paper uses the timing differences between the COVID-19 national emergency, school closures, and stay-at-home order to determine the impact that each of these events had on the decline in child maltreatment referrals for the full year. Prior research has looked at either school closures or stay-at-home orders in isolation and only examined the effect from March to May (e.g. Baron et al, 2020; Bullinger et al., 2020). This is the first study to provide the impacts of all three events for the full year. To differentiate these three impacts, separate regression equations are estimated with an independent variable equal to the proportion of the quarter in which the event happened. I find evidence that the largest decline in reporting came when the three events were happening concurrently. Alternatively, the pandemic, without school closures and stay-at-home orders, had the smallest impact on underreporting. Understanding the difference in reported maltreatment as a result of the pandemic and policy responses to curb the spread of the virus helps quantify the effectiveness of these policy responses. In addition, understanding the difference in reported maltreatment as a result of the pandemic and school

[2] Brown & De Cao (2020) find that unemployment is positively correlated with child maltreatment, Griffith (2020) explains how limited availability of social supports and child care can lead to parental burnout, which in turn can result in neglect and abuse (Mikolajczak et al., 2019), and the WHO published a brief explaining the links between alcohol abuse and neglect (WHO, nd).

closures contributes to the emerging body of literature emphasizing teachers' roles in detecting child maltreatment (Fitzpatrick et al., 2020; Cabrera-Hernandez & Padilla-Romo, 2020).

The last contribution of this paper is to identify the impact of the pandemic and pandemic-induced policies on the *type* of child maltreatment reported and substantiated. Child welfare experts observed an increase in serious abuse (Hofmann, 2021), and doctors claimed the severity of the abuse they saw in the ER at the start of the pandemic was much worse (Schmidt & Natanson, 2020). In Colorado, I do not find evidence of these claims. Overall, the proportion of neglect, physical abuse, and sexual abuse allegations is unchanged from 2019 to 2020. In addition, I use economic and seasonal trends to predict the type of maltreatment that might be going unreported. Based on this approach, victims of neglect are most likely being missed. In order to better prepare and target interventions, it is important to understand the type of maltreatment occurring and being underreported as a result of the COVID-19 pandemic.

2. The COVID-19 Pandemic and Child Maltreatment in Colorado

The COVID-19 pandemic national emergency was announced on March 13, 2020 and continued through the end of the year. To curb the spread of the virus in the early months, schools halted in-person learning, stay-at-home orders were issued, and non-essential employees worked from home. The unemployment rate rose to an all-time high of 14.8 percent in April 2020 and remained above 6 percent for the remainder of the year (Trading Economics, nd). Frequency of alcohol consumption increased by 14 percent (Pollard et al., 2020), domestic violence calls increased by 7.5 percent (Leslie & Wilson, 2020), people's mental health deteriorated (Brodeur et al., 2020), and parental burnout probably increased (Griffith, 2020). Despite these hardships and risk factors of child maltreatment, hotline calls to state agencies plummeted (Schmidt & Natanson, 2020), raising concerns that abuse and neglect are going unreported (MacFarlane et al., 2020).[3] Research suggests that these drops in reporting came from the pandemic-induced school closures which limited interactions with mandatory reports (Baron et al., 2020, Cabrera-Hernández & Padilla-Romo, 2020).

Colorado is no exception to the situation described above. School closures began at the end of March and continued until September. Compared to non-pandemic years, children were out of

[3] Alternatively, Ortiz et al. (2021) find that the volume of text messages to Childhelp, the only national hotline providing counseling services with a focus on child abuse and neglect, increased in 2020 compared to 2019.

Covid Economics 82, 23 June 2021: 10-48

school for three additional months, but they continued to have access to school meals (Grewe, 2020). In addition, Colorado issued a stay-at-home order from March 26 to April 26. Colorado permitted going outside during the stay-at-home order as long as social distancing was followed. In fact, the public health order specifically listed walking, hiking, skiing, snowshoeing, biking, and running as acceptable activities (Grewe, 2020). People could also go to the grocery store, liquor store, convenience store, cannabis store, banks, and pharmacies (Grewe, 2020). The unemployment rate fluctuated from 6.8 to 11 between April to December and alcohol sales increased by 8 percent.[4]

Figure 1 shows the number of child maltreatment referrals received and screened-in for investigation from 2006 through 2020 in the state. From 2006 to 2020, there was a steady incline, with a steeper incline following the introduction of the statewide hotline in 2015. Between April and June of 2020, the number of total referrals and reports screened-in[5] dropped by 31 and 26 percent, respectively, relative to the same time period in 2019. In the remaining months of the year, child maltreatment reporting rebounded somewhat, but referrals and reports still remained below pre-pandemic levels. The bottom graph of Figure 1 shows the number of allegations reported by maltreatment type, indicating the biggest drop and rebound in neglect allegations, and a small uptick in sexual abuse allegations. According to the Colorado Department of Human Resources, calls from education and medical personnel decreased by 30 and 11 percent, respectively; however, calls from friends and family increased by 5 percent (CDHS, 2021).

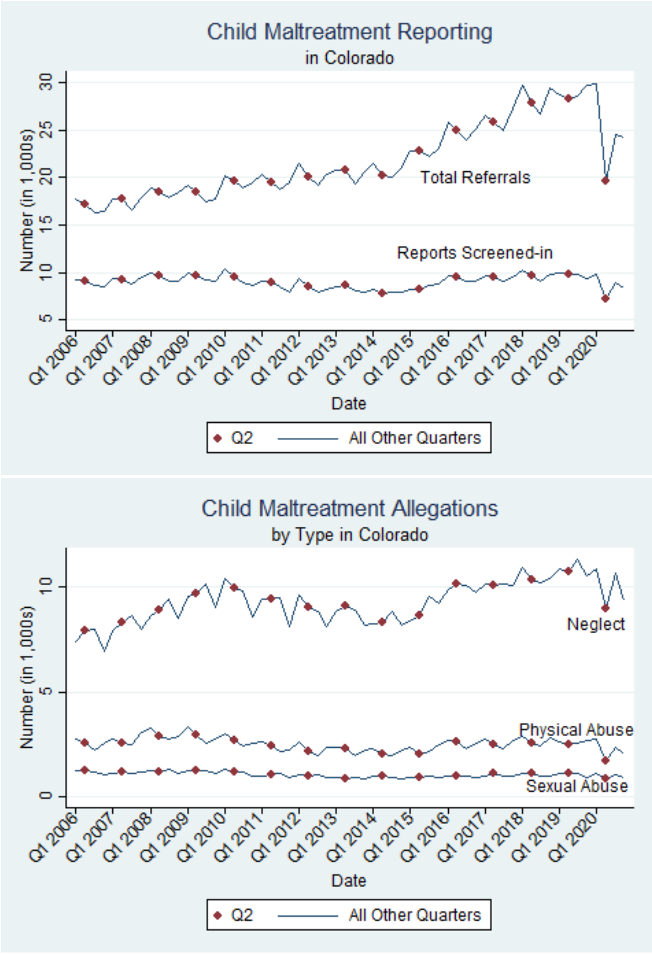
Referrals of child maltreatment are a function of actual incidences and reporting. The current pandemic, which increased economic hardship while reducing contact with mandatory reporters, poses a particularly unique challenge for child welfare agencies to detect child maltreatment. These two opposing forces will attenuate the impact of the pandemic on child maltreatment reporting towards zero. Alternatively, school closures and the stay-at-home order limited interactions with mandatory reporters. These two events should be driving the decline in reporting, and might explain some of the rebound in reporting after they ended.

[4] Author's calculations based on unemployment data from the BLS and alcohol sales data from NIAAA. These data sources are described in the next section. Additionally, the Liquor Excise Tax Reports show a similar increase and be accessed through Colorado's Department of Revenue.

[5] Total referrals include both the calls to the hotline that are screened out and in. Referrals that are screened-in are also referred to as reports. There is no additional follow-up for referrals that are screened-out, but reports are investigated for child maltreatment.

CEPR PRESS

Figure 1: Child Maltreatment Referrals in Colorado from 2006 to 2020



Notes: This figure shows the trend in child maltreatment reporting from 2006 to 2020. The top graph shows the total number of child maltreatment referrals (in thousands) reported to child welfare agencies in the state as well as the number of reports screened-in (in thousands). The bottom graph shows the number of allegations by maltreatment type (in thousands). The total number of maltreatment allegations in the bottom graph is less than the total referrals because not all referrals are screened-in and investigated. In addition, the total number of maltreatment allegations is greater than the screened-in reports because a report can be assigned multiple maltreatment types.

Covid Economics 82, 23 June 2021: 10-48

Since the stay-at-home order limited all potential interactions with mandatory reporters and occurred concurrently with the first month of school closures, we should see this event driving the decline experienced between April and June. While the stay-at-home order limited interactions, it also may have increased household stress and parental burnout, potentially more so for those who complied. Parental burnout can manifest into neglect (Mikolajczak et al., 2019). As a result, we expect to see differential effects in child maltreatment for counties with higher compliance relative to counties with lower compliance. This behavior might also be able to explain the uptick in neglect and sexual abuse.

3. Data

The data for this study come from multiple public sources. The child maltreatment data come from Colorado's Department of Human Services (CDHS), which provides real-time quarterly counts of calls made to the child abuse and neglect hotline for each county in Colorado starting in 2006. For this time-sensitive project, the CDHS data are preferred over the National Child Abuse and Neglect Data System (NCANDS) because the national level data have a two-year time lag and only provide screened-in reports in counties with more than 1,000 records, whereas the real-time CDHS data provide the total number of hotline calls at the county level. Another advantage of the CDHS data, relative to other states' real-time data, is that they provide the type of alleged maltreatment and the finding of the allegation. These data are used to test the hypothesis that child welfare experts have posited; more severe cases of abuse will result from the pandemic. Prior research in Florida and New York did not estimate the real-time composition of child maltreatment reports (Baron et al., 2020; Rapaport et al., 2020), and research in Indiana found an increase in neglect, not physical abuse (Bullinger et al., 2020).

One drawback of these data, is that the analysis is limited to a single state. The extent to which these results can be generalized to the entire country is questionable. Colorado had one of the highest child maltreatment referral rates of 85.2 referrals per 1,000 children in 2019 (ACF, 2021). The average referral rate for states across the country was 59.5 (ACF, 2021). In addition, Colorado screened out more referrals than the average state. Colorado screened out 66.4 percent of their referrals in 2019, whereas the average screen-out rate was 40.7 percent (ACF, 2021). Of the calls that were screened-in, about 34 percent were substantiated in Colorado, compared to an average of 29 percent across the country (ACF, 2021). Finally, the most common types of maltreatment in both the US and Colorado are neglect, abuse, and sexual abuse (ACF, 2021); however, neglect is

Covid Economics 82, 23 June 2021: 10-48

relatively higher and physical abuse is relatively lower in Colorado compared to the typical state. While Colorado may not be representative of the typical state in the US, these results are essential to provide more evidence of the impacts of the pandemic and pandemic-induced policies.

The remaining data come from multiple sources and are used to supplement the main analyses. First, I use employment and population data from the Bureau of Labor Statistics (BLS) and US Census to control for changes in economic conditions. The BLS provides county and state-level unemployment rates and employment counts, quarterly from 2006 through 2020, and the Census provides the county population size, annually from 2008 to 2019. To estimate the 2020 population numbers, I use the 3-year average percent change in each county.[6] The population size and employment counts are used to determine the employment to population ratio. In addition, I use the population to determine county-level child maltreatment rates. Next, I use alcohol sales data from the "Surveillance Report #115" and "Alcohol Sales during the COVID-19 Pandemic" files, maintained by National Institute on Alcohol Abuse and Alcoholism,[7] to proxy for alcohol consumption at the state-level. This estimate, in combination with the unemployment rate, is used to create a second counterfactual maltreatment number that accounts for an economic hardship and potential coping strategy. Finally, to proxy for stay-at-home order compliance, I obtain county-level data on COVID-19 cases and deaths from Colorado's Outbreak Data, maintained by Colorado's Department of Public Health and Environment (CDPHE).[8] These data are updated weekly and available online for transparency and evidenced-based decision-making, but may not be comparable across counties over time and should not be used to associate exposure risk with certain settings (CDPHE, nd). I use these data from the beginning of the pandemic (March 14, 2020 to May 10, 2020) to observe how caseloads changed *within* a county prior to and during the stay-at-home order to get an idea of stay-at-home order compliance. While these four additional sources of data do not control for all potential confounders, they enrich analyses that solely rely on seasonal and longitudinal trends.

Appendix Table 1 provides statewide differences in child maltreatment reporting, economic conditions, and alcohol sales for each quarter between 2019 and 2020 in Colorado. In addition,

[6] More specifically, I first calculated the percent changes in population size from 2016 to 2017, 2017 to 2018, and 2018 to 2019. Then, I calculated the 3-year average and used this average to estimate the 2020 population size.
[7] These data can be found here.
[8] These data were downloaded March 24, 2021 from https://covid19.colorado.gov/covid19-outbreak-data.

percent changes are provided. Overall, child maltreatment reporting declined by 15 percent, with the biggest decline of 31 percent occurring between April and June. The proportion of screened-in and substantiated reports remained similar between 2019 and 2020. Appendix Table 2 provides summary statistics of child maltreatment reporting and economic conditions for all 64 counties over the 4 quarters and 13 years. The average number of referrals received in a county during a given quarter between the years 2008 and 2020 is 18, per 1,000 children. I also provide the 2019 and 2020 averages and a p-value indicating if they are statistically different from each other.

4. Empirical Strategy

Similar to Baron et al. (2020), I first predict the counterfactual number of child maltreatment referrals, screened-in reports, and substantiated reports for the state of Colorado by estimating the following equation:

$$Y_{qy} = \beta_0 + \varphi_q + f_g(qy) + \varepsilon_{qy} \qquad (1)$$

Where Y is the outcome of interest (i.e. number of referrals made to the hotline, number of reports screened-in, number of substantiated reports, etc.) in Colorado during quarter q of year y, φ_q is the quarter fixed effect included to capture seasonal trends,[9] $f_g(qy)$ is a polynomial in time of order g, and ε_{qy} is the error term. In the main specification the polynomial takes a cubic form; however, the counterfactual results are similar across alternative specifications.[10] This equation is estimated for each of the four quarters from the years 2006 to 2019. These estimates are then used to predict the outcomes for each quarter in year 2020. This approach assumes that the number of maltreatment referrals and reports would have been similar in 2020 as 2019, had the pandemic not occurred. Alternatively, the hardships and stresses brought about by the pandemic might increase child abuse and neglect. In attempt to capture the increase in maltreatment due to hardships, I estimate equation 1 again controlling for the unemployment rate and alcohol purchases. This approach assumes the relationships between unemployment and child maltreatment and alcohol purchases and child maltreatment are similar in 2020 and 2019. Estimating two counterfactuals based on seasonal and longitudinal trends is useful as there is no feasible control group since the

[9] Quarters one and four experience the highest call volume, whereas the quarters spanning the summer experience the lowest call volumes. This can be seen in Figure 1.

[10] See Appendix Figure 1 for a comparison of the different approaches that use a linear and quadratic polynomial.

announcement of the national emergency and subsequently policy responses occurred at the same time for the entire country.

After understanding the difference between the counterfactual and actual scenarios, the next step is to understand how much of these differences are driven by the pandemic, the pandemic-induced school closures, and the pandemic-induced stay-at-home order. I estimate the following equations to differentiate these three effects:

$$Y_{cqy} = \beta_0 + \beta_1 covid_{\,qy} + \varphi_q + \gamma_y + \rho_c + \varepsilon_{cqy} \qquad (2)$$

$$Y_{cqy} = \alpha_0 + \alpha_1 schclo_{\,qy} + \varphi_q + \gamma_y + \rho_c + \varepsilon_{cqy} \qquad (3)$$

$$Y_{cqy} = \delta_0 + \delta_1 sah_{\,qy} + \varphi_q + \gamma_y + \rho_c + \varepsilon_{cqy} \qquad (4)$$

Where Y is the outcome of interest in county c during quarter q of year y, φ_q, γ_y, and ρ_c are the quarter, year, and county fixed effects, respectively, and ε_{cqy} is the error term. The independent variables of interest, *covid, schclo*, and *sah,* identify the proportion of the quarter in which the condition exists in quarter q of year y. For example, the COVID-19 pandemic national emergency was announced March 13, 2020 and continued through 2020,[11] so *covid* is assigned a value of one-sixth in quarter 1 in 2020, and a value of one for the remaining quarters in 2020. The national emergency forced schools to close in March and delayed openings, so *schclo* equals one-sixth in quarter one, two-thirds in quarter two, and one-sixth in quarter three during 2020. Lastly, in attempt to slow the spread of the virus, Colorado issued a stay-at-home order from March 26th to April 26th, so *sah* in equation 4 is assigned one-third in quarter two of year 2020 and zero otherwise.[12] Table 1 provides the dates and values defined and Figure 2 provides a graphic representation of the timing of the events.

[11] See The White House notice on the continuation of the National Emergency found here.
[12] In all equations, standard errors are clustered at the county by quarter level. Results are similar when standard errors are clustered at the county level and available upon request.

Table 1: Timeline of Events and Independent Variable Values

Event	Dates	Independent Variable Values
COVID-19 National Emergency	March 13, 2020 – March 2021[13]	$covid_{qy} = \begin{cases} 0 \ if \ y < 2020 \\ \frac{0.5}{3} if \ y = 2020 \cap q = 1 \\ 1 \ if \ y = 2020 \ \cap q > 1 \end{cases}$
School Closures in Colorado	March 16, 2020 – August 24, 2020[14]	$schclo_{qy} = \begin{cases} \frac{0.5}{3} if \ y = 2020 \cap q = 1 \\ \frac{2}{3} if \ y = 2020 \cap q = 2 \\ \frac{0.5}{3} if \ y = 2020 \ \cap q = 3 \\ 0 \ otherwise \end{cases}$
Stay-at-home Order in Colorado	March 26, 2020 – April 26, 2020[15]	$sah_{qy} = \begin{cases} \frac{1}{3} if \ y = 2020 \cap q = 2 \\ 0 \ otherwise \end{cases}$

Notes: This table lists the dates of the COVID-19 national emergency, stay-at-home order, and school closures in Colorado. Using these dates, the independent variables are defined. The variable y indicates the year, and the variable q indicates the quarter. The independent variable is rounded to the nearest half month, out of 3 months. For example, the COVID-19 pandemic was announced as a national emergency March 13th, so about 0.5 months out of 3 were impacted by the pandemic. The national emergency existed for the rest of 2020, so for the remaining three quarters, three out of three months were impacted, which equals one. The stay-at-home order primarily took place in April, so in 2020 for quarter 2, *sah* equals one-third and zero otherwise. Lastly, school closure is defined based on the month impacted by the pandemic. For example, quarter two consists of April, May, and June, and in June, schools would have been closed regardless of the pandemic, so *schclo* is two-thirds, and not three-thirds (i.e. one).

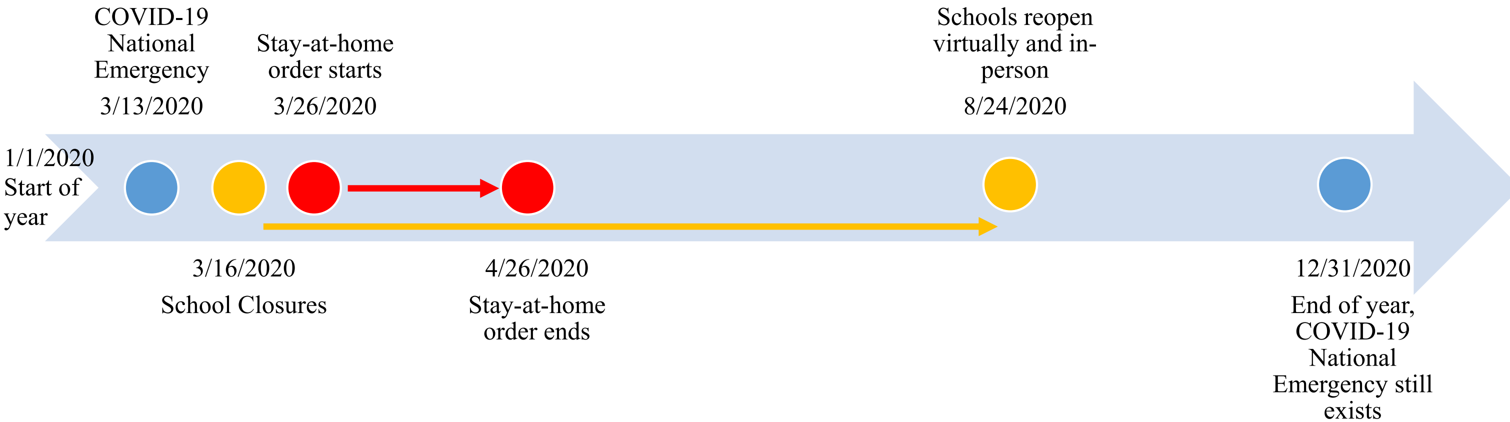
[13] See The White House notice on the continuation of the National Emergency found here.

[14] Between the following two sources, https://co.chalkbeat.org/2020/3/12/21178764/the-complete-list-of-coronavirus-related-colorado-school-closures and https://www.denverpost.com/2020/07/01/colorado-schools-reopening-coronavirus-covid/, most school districts in Colorado closed on March 16, 2020 and most districts delayed opening in the fall by a few weeks, resulting in opening dates between August 24 and September 1. In-person and virtual learning varied by district, but since Colorado counties and school districts do not align, the school closure variable is the same for all counties, regardless of instructional mode and based on the date that impacts most of the state. As a robustness check, I allow the opening dates to vary for a few counties that are clearly defined, such as those in the Denver metro-area, but the results are similar.

[15] See https://www.westword.com/news/covid-19-colorado-stay-at-home-order-shorter-than-most-in-america-11682795 for a list of state closing and opening dates. Some jurisdictions, like Denver, extended their stay-at-home order, and Colorado issued a "safer-at-home" order following the stay-at-home order. See https://www.kktv.com/content/news/Gov-Polis-issues-Executive-Order-on-Safer-at-Home-569966341.html for more details. These variations are not accounted for in this analysis.

Covid Economics 82, 23 June 2021: 10-48

Figure 2: Timeline of Events in 2020



Notes: This figure plots the timeline of events during 2020 in Colorado. The COVID-19 National Emergency was announced on March 13, 2020 and continued through the year. Schools closed on March 16 and the stay-at-home order began March 26. The stay-at-home order ended a month later, and schools reopened for in-person and virtual learning at the end of August. The red arrow shows when the stay-at-home order happened, the yellow arrow shows when schools were closed for in-person learning, and the blue arrow shows when the pandemic existed. All three events happened concurrently from March 26 to April 26, and two of the events happened concurrently from April 26 to August 24. After August 24, only the COVID-19 national emergency was happening.

These three effects cannot be estimated together because they are correlated with each other due to the timing of the events. For example, when the stay-at-home order is in effect, schools are closed and the pandemic exists. After the stay-at-home order is lifted, when schools are closed, the pandemic exists. The timing overlap of these events implies that $\delta_1$captures the impact of the pandemic, school closures, and stay-at-home order concurrently on child maltreatment reporting. Similarly, α_1 captures the impact of the pandemic-induced school closures, and β_1 estimates the impact of the pandemic without stay-at-home orders or school closings, like the end of 2020. This setup implies $\delta_1 > \alpha_1 > \beta_1$.

Finally, to determine differential effects of the stay-at-home order by compliance, I estimate the following equation:

$$Y_{cqy} = \delta_0 + \delta_1 sah_{\ qy} + \delta_2 cc_c + \delta_3 sah_{\ qy}\ x\ cc_c + \varphi_q + \gamma_y + \rho_c + \varepsilon_{cqy} \qquad (5)$$

Where the majority of the terms are defined above. *cc* measures the pre-stay-at-home order COVID-19 cases per 100,000 residents for county c, so δ_2 estimates the relation between COVID-19 cases and child maltreatment reporting, and $\delta_3$estimates the interaction effect between the stay-at-home order and COVID-19 cases on child maltreatment. A positive coefficient on δ_3 means the stay-at-home order increased child maltreatment reporting for counties with higher COVID-19 cases, relative to counties with no COVID-19 cases as of March 26th.

The validity of this approach to yield causal estimates relies on two assumptions. First, counties with COVID-19 cases prior to March 26th had similar child maltreatment reporting trends as counties with no COVID-19 cases prior to March 26th. Second, counties with COVID-19 cases prior to March 26th were more compliant to the stay-at-home order. One crude proxy for compliance is the number of COVID-19 cases during the stay-at-home order. Counties with fewer cases per 100,000 residents are considered more compliant, especially if they had cases prior to the stay-at-home order. Appendix Table 3 provides summary statistics of child maltreatment reporting by COVID-19 cases, and Appendix Figure 2 plots the relationship between COVID-19 cases prior to the stay-at-home order versus during the stay-at-home order with the regression adjusted correlation coefficient. Neither of these set of results provide convincing evidence that the two assumptions are satisfied, so this analysis testing the differential effects by compliance is exploratory.

To understand the direction of the potential bias, I estimate the relationship between compliance on pre-pandemic child maltreatment. For this exercise, compliance is measured as the

Covid Economics 82, 23 June 2021: 10-48

ratio between COVID-19 cases prior to and during the stay-at-home order.[16] In this setup, counties with ratios greater than or equal to one are considered more compliant than counties with ratios less than one. I do not find strong evidence that compliance is correlated with child maltreatment prior to the pandemic, thus I cannot sign the potential bias.

5. Results

5.1. Counterfactual number of referrals and reports

Figure 3 shows the predicted versus actual number of referrals from 2006 to 2020. In 2019, there were a total of 115,178 referrals made, 38,950 reports were screened-in, and 13,730 reports were substantiated. In contrast, in 2020, there were a total of 98,158 referrals made, 34,121 reports were screened-in, and 12,684 reports were substantiated. Counterfactual 1 shows the predicted number of referrals assuming the pandemic had not occurred, and counterfactual 2 shows the predicted number of referrals after accounting for increases in unemployment and alcohol sales, de facto consumption.

Comparing the actual number of referrals to counterfactual 1, an estimated 30,276 referrals went unreported in 2020. Alternatively, recognizing that the pandemic has brought on significant hardships, we might expect the number of children suffering from maltreatment to be even greater in 2020 relative to 2019. Comparing the actual number of referrals to counterfactual 2, an estimated 38,794 referrals went unreported in 2020.

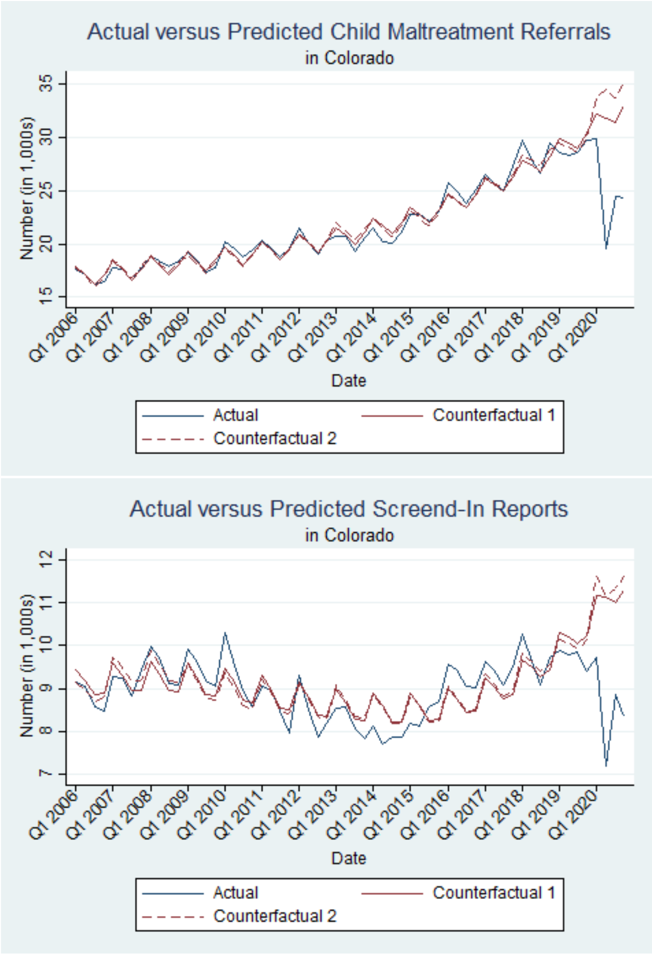
Colorado has a high referral rate, but a high proportion of referrals are screened-out and unsubstantiated. Next, I compare the actual screened-in reports to the predicted numbers of screened-in reports to investigate whether the screening process changed during the pandemic. If the screening process remained the same, then the proportion of the predicted screened-in reports would be the same as the proportion of the actual screened-in reports. Figure 3 shows the predicted versus actual number of screened-in reports from 2006 to 2020. Between 10,500 to 11,500 fewer reports were screened-in in 2020 compared to counterfactual estimates. The proportion of reports that were screened-in in 2020 and the proportion of reports that should have been screened-in

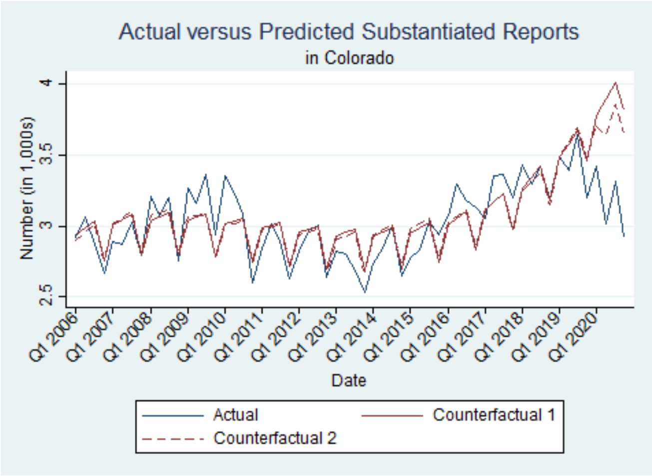
[16] The majority of counties (45) reported zero COVID-19 cases prior to and during the stay-at-home order. For these counties the ratio is set equal to one. In the remaining 19 counties, there were COVID-19 cases before and/or during the stay-at-home order. In one county, they reported zero COVID-19 cases during the stay-at-home order, resulting in an invalid ratio (i.e. zero in the denominator). For this county, I set the ratio equal to the number of cases prior to the stay-at-home order.

Covid Economics 82, 23 June 2021: 10-48

based on the counterfactual estimates was 33 to 35 percent, indicating the screening process remained the same during the pandemic. Lastly, I compare the actual number of substantiated reports to the predicted number of substantiated reports to determine whether the nature of child maltreatment changed during the pandemic. Figure 3 shows the predicted versus actual number of substantiated reports from 2006 to 2020; 2,200 to 2,800 substantiated reports were missed. The different trends for each line and the uptick in actual reports in quarter 3 of year 2020 make it difficult to interpret how substantiated maltreatment has changed over the year.

Figure 3: Actual versus Predicted Child Maltreatment Referrals and Reports in Colorado from 2006 to 2020



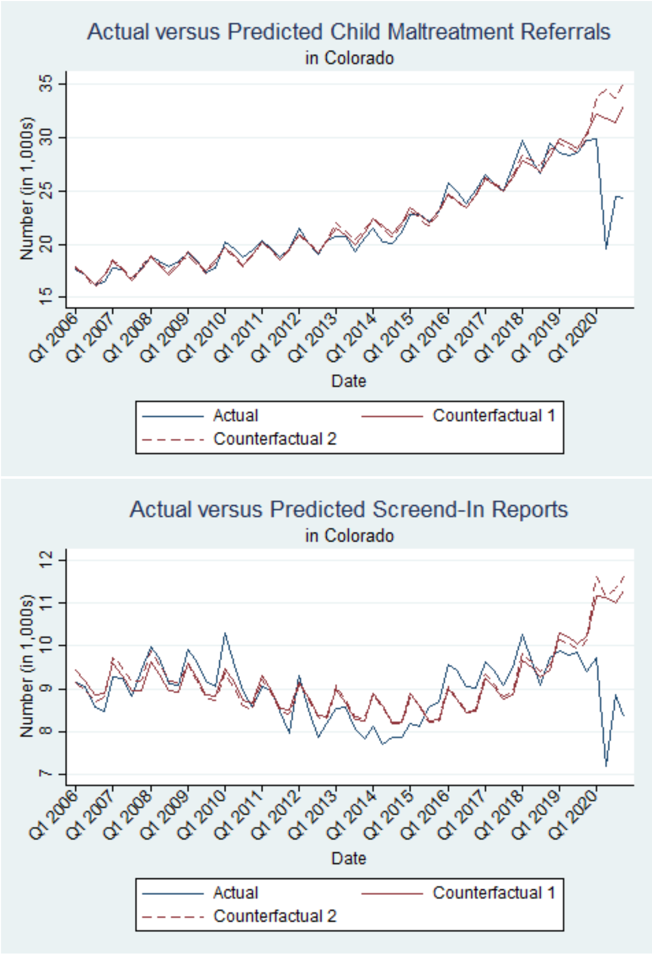


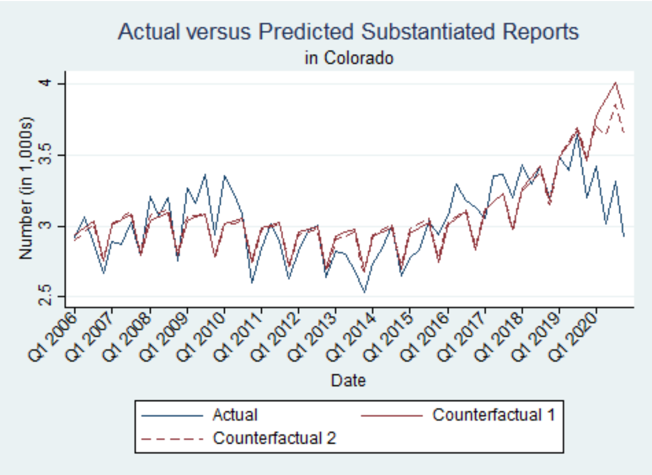
Notes: These figures plot the actual versus predicted counts of child maltreatment referrals (top graph), screened-in reports (middle graph), and substantiated reports (bottom graph). Two counterfactuals are estimated: counterfactual one assumes child maltreatment would have been the same in 2020 as 2019 had the pandemic not occurred, and counterfactual two accounts for changes in maltreatment as a result of changes in the unemployment rate and alcohol purchases.

In addition to estimating counterfactuals for the screened-in and substantiated reports, the CDHS data also allow me to estimate counterfactuals for the composition of substantiated reports by maltreatment type. Figure 4 plots the predicted versus actual number of substantiated neglect, physical abuse, and sexual abuse allegations. The majority of unreported victims seem to be suffering from neglect.

Covid Economics 82, 23 June 2021: 10-48

Figure 4: Actual versus Predicted Child Maltreatment Referrals and Reports in Colorado from 2006 to 2020





Notes: These figures plot the actual versus predicted counts of child maltreatment referrals (top graph), screened-in reports (middle graph), and substantiated reports (bottom graph). Two counterfactuals are estimated: counterfactual one assumes child maltreatment would have been the same in 2020 as 2019 had the pandemic not occurred, and counterfactual two accounts for changes in maltreatment as a result of changes in the unemployment rate and alcohol purchases.

5.2. Impact of the COVID-19 pandemic, school closures, and the stay-at-home order on child maltreatment reporting

Table 2 provides the main results of the paper. All else equal, an additional quarter with the COVID-19 pandemic reduced the number of referrals made to the hotline by 2.5 per 1,000 children (or 10% relative to the average 2019 referral rate in a county). The screened-in report rate and substantiation rate are not statistically different as a result of the COVID-19 pandemic. All else equal, an additional quarter with pandemic-induced school closures reduced the number of referrals by 7.9 per 1,000 children (or 32% relative to the average 2019 referral rate in a county) and reports screened-in by 1.8 per 1,000 children (or 24% relative to the average 2019 report rate in a county). Finally, all else equal, an additional quarter with a stay-at-home order reduced the number of referrals by 14.8 per 1,000 children (or 60% relative to the average 2019 referral rate in a county) and reports screened-in by 3.3 per 1,000 children (or 42% relative to the average 2019 report rate in a county). Similar to the COVID-19 pandemic, neither school closings nor the stay-at home order had a statistically significant impact the substantiation rate.

Covid Economics 82, 23 June 2021: 10-48

Table 2: Estimated Impacts of COVID-19 Pandemic, School Closures, and Stay-at-home Order on Child Maltreatment Reporting in Colorado

	Main Outcomes			Type of Maltreatment		
	Total Referrals (per 1,000 children)	Reports Screened-in (per 1,000 children)	Substantiated Reports (per 1,000 children)	Neglect Allegations (per 1,000 children)	Physical Abuse Allegations (per 1,000 children)	Sexual Abuse Allegations (per 1,000 children)
Ind. Var.: COVID-19	-2.541*	-0.079	-0.876	-0.445	-0.084	0.251
	(1.415)	(0.710)	(0.715)	(1.078)	(0.279)	(0.178)
Ind. Var.: School Closure	-7.874***	-1.838**	0.615	-0.961	-0.229	0.115
	(1.829)	(0.846)	(0.966)	(1.239)	(0.512)	(0.350)
Ind. Var.: Stay-at-home Order	-14.787***	-3.273**	0.628	-2.430	-0.638	0.265
	(3.070)	(1.479)	(1.764)	(2.260)	(0.909)	(0.625)
2019 Average	24.66	7.83	2.61	8.72	2.2	0.79
Observations	3,328	3,328	3,328	3,328	3,328	3,328

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors, clustered at the county-by-quarter level, in parentheses. Each column indicates an outcome of interest, provided as a rate per 1,000 children. Each row represents a separate regression analysis, so row 1 reports the coefficient from equation 2 where *covid* is the independent variable of interest. Row 2 reports the coefficient from equation 3 where *schclo* is the independent variable of interest, and row 3 reports the coefficient from equation 4 where *sah* is the independent variable of interest. Each regression includes year, county, and quarter fixed effects.

Rescaling the quarterly effect to a monthly effect, implies that an additional month of the COVID-19 pandemic, school closures, and stay-at home order reduced child maltreatment reporting by 0.85, 2.6, and 4.9 referrals per 1,000 children, respectively. The effect of the stay-at-home order is almost six times as large as the effect of the pandemic and almost twice as large as the effect of the school closings. Finally, rescaling the monthly impact to an annual impact based on the number of months for each of the events implies the COVID-19 pandemic, school closures, and stay-at-home order reduced maltreatment reporting by 8, 7.9, and 4.9 referrals per 1,000 children, respectively. With a child population of 1.26 million, approximately 10,000, 9,900, and 6,200 referrals went unreported as a result of the COVID-19 pandemic, school closures, and stay-at-home order, respectively, in Colorado.

5.3. Changes in type of maltreatment

So far this paper has demonstrated that the COVID-19 pandemic, and subsequent school closures and stay-at-home order drastically decreased the child maltreatment referral and report rate, but had no statistically significant impact on the substantiation rate. Next, I explore whether the type of maltreatment reported changed during the pandemic. The right side of Table 2 provides the results from estimating equations 2-4 for the neglect, physical abuse, and sexual abuse referral rates. Overall, there is no evidence that the pandemic altered the type of maltreatment reported. The direction of the coefficients implies fewer neglect and physical abuse allegations, and more sexual abuse allegations were reported as a result of the pandemic and subsequent policy responses; however, none of these estimates are statistically significant. As a result of statistically insignificant changes in the type of maltreatment referrals made, there are no statistically significant changes in the type of maltreatment substantiated.[17]

5.4. Relationship between stay-at-home order compliance and child maltreatment reporting

Finally, Table 3 provides the estimates from the interaction between the stay-at-home order and COVID-19 cases. All else equal, an additional quarter under the stay-at-home order in counties with COVID-19 cases is associated with a smaller decline in the number of total referrals and screened-in referrals by 0.09 and 0.05 per 1,000 children, relative to counties without COVID-19 cases. Moreover, this interaction analysis suggests that the stay-at-home order is correlated with

[17] Results available upon request.

Covid Economics 82, 23 June 2021: 10-48

counties with COVID-19 cases experiencing a smaller decline in the neglect allegation rate and an increase in the sexual abuse allegation rate, relative to counties without COVID-19 cases.

Table 3: Estimated Impact of Stay-at-home Order Interacted with COVID-19 Cases on Child Maltreatment Reporting

	Main Outcomes			Type of Maltreatment		
	Total Referrals (per 1,000 children)	Reports Screened-in (per 1,000 children)	Substantiated Reports (per 1,000 children)	Neglect Allegations (per 1,000 children)	Physical Abuse Allegations (per 1,000 children)	Sexual Abuse Allegations (per 1,000 children)
Stay-at-home	-15.413***	-3.607**	0.664	-2.879	-0.619	0.220
	(3.104)	(1.491)	(1.812)	(2.293)	(0.935)	(0.642)
COVID-19 Cases	0.024	-0.123***	0.045***	-0.030	-0.037	-0.014
	(0.077)	(0.030)	(0.015)	(0.051)	(0.023)	(0.012)
Stay-at-home x COVID-19 Cases	0.094***	0.050***	-0.005	0.067***	-0.003	0.007**
	(0.015)	(0.006)	(0.008)	(0.011)	(0.005)	(0.003)
Observations	3,328	3,328	3,328	3,328	3,328	3,328

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors, clustered at the county-by-quarter level, in parentheses. Each column indicates an outcome of interest, provided as a rate per 1,000 children. Row 1 reports the coefficient on the stay-at-home order, row 2 reports the coefficient on the number of pre-stay-at-home order COVID-19 cases, and row 3 reports the coefficient on the interaction term between the stay-at-home order and pre-stay-at-home COVID-19 cases. Each regression includes year, county, and quarter fixed effects.

5.5. Alternative analyses and permutation tests

To determine the sensitivity of the main results, I estimate alternative analyses varying the sample size and control variables. Table 4 reports the results for different measures of the child maltreatment referral rate across different analyses. Column 1 provides the main results again. Columns 2 and 3 provide results from varying the sample size, and column 4 provides results that include county-level controls for economic conditions. The first panel of results shows the change in the rate, per 1,000 children, the second panel uses logged values, and the third panel uses the level values.

Overall, the impact of the pandemic, school closures, and stay-at-home order on the referral rate (panel 1) and logged number of referrals (panel 2) is similar across varying sample sizes. In column 2, the time period is restricted to the years 2010 to 2020, to exclude any impacts of the Great Recession. In column 3, the Denver metro-area is excluded to test whether these results are generalizable to all counties in Colorado or unique to the most populous areas. When observing total referral levels (panel 3), the coefficients are not sensitive to excluding the Great Recession years, but they are drastically reduced when excluding the Denver metro-area. This analysis indicates that relatively more referrals are going unreported in the Denver metro-area relative to other parts of the state, which makes sense since there are more people and children in the Denver metro-area. This sensitivity also underscores the importance of using rates, and not levels. In all cases, including the economic conditions reduces the magnitude of the effect size, relative to the main specification. However, the change in magnitude is not statistically different from the main estimates. The same conclusions apply across analyses variations for the screened-in and substantiation rates.[18]

[18] See Appendix Table 4 and 5 for the results.

Table 4: Robustness Analyses for Child Maltreatment Referrals

		(1) Main Results	(2) Post 2010	(3) Exclude Denver Metro-area	(4) Include Economic Controls
Panel A: Rate (per 1,000 children)	COVID-19	-2.541*	-2.514*	-1.868	-1.053
		(1.415)	(1.418)	(1.632)	(1.547)
	School Closure	-7.874***	-7.358***	-7.684***	-6.402***
		(1.829)	(1.799)	(2.107)	(1.917)
	Stay-at-home	-14.787***	-14.019***	-14.396***	-12.415***
		(3.070)	(3.056)	(3.538)	(3.274)
Panel B: Log	COVID-19	-0.133*	-0.131*	-0.096	-0.056
		(0.074)	(0.072)	(0.085)	(0.080)
	School Closure	-0.420***	-0.392***	-0.397***	-0.353***
		(0.114)	(0.111)	(0.131)	(0.120)
	Stay-at-home	-0.780***	-0.744***	-0.733***	-0.681***
		(0.192)	(0.188)	(0.221)	(0.206)
Panel C: Levels	COVID-19	-110.255***	-110.765***	-42.456*	-73.236**
		(34.192)	(29.319)	(24.229)	(32.786)
	School Closure	-146.414***	-141.809***	-72.856***	-89.796***
		(32.469)	(36.791)	(25.657)	(28.643)
	Stay-at-home	-301.492***	-295.858***	-148.273***	-201.033***
		(61.579)	(67.786)	(48.256)	(53.836)
	Observations	3,328	2,560	2,860	3,328

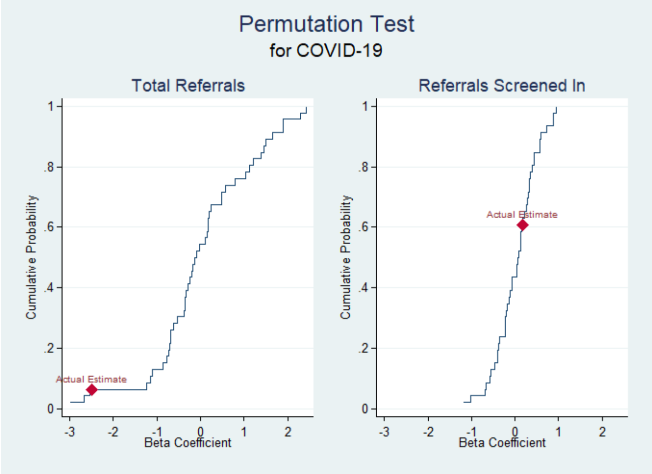
Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors, clustered at the county-by-quarter level, in parentheses. Each column indicates a separate regression analysis. The first column provides the main results. The second column is post-2010. The third column excludes the Denver metro-area, and the fourth column includes controls for economic conditions (i.e. the unemployment rate and employment-to-population ratio). Each row represents a separate regression analysis for the three independent variables of interest, so row 1 reports the coefficient from equation 2 where *covid* is the independent variable of interest. Row 2 reports the coefficient from equation 3 where *schclo* is the independent variable of interest, and row 3 reports the coefficient from equation 4 where *sah* is the independent variable of interest. There are 3 separate panels as well. Panel one reports the effect on the total referral rate, per 1,000 children. The second panel reports the effect on the log total referrals, and the third panel reports the effect on the total referrals (level). Each regression includes year, county, and quarter fixed effects.

Covid Economics 82, 23 June 2021: 10-48

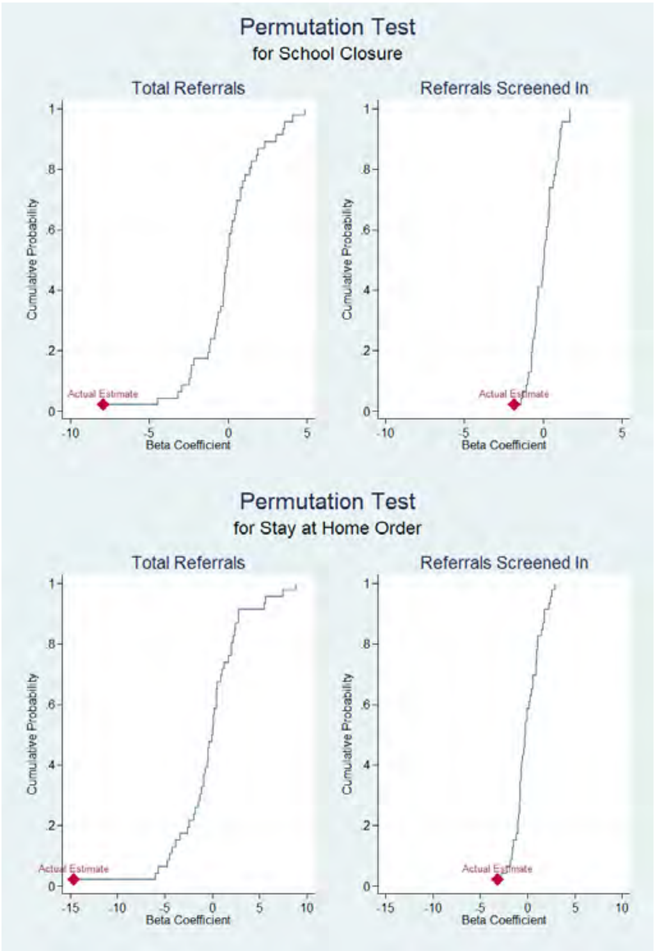
Since the pandemic started in 2020, there are only a few treated observations. Few treated observations can lead to improper inference (Cameron et al., 2008; MacKinnon & Webb, 2017, 2018; Ferman & Pinto, 2019). One way to correct for this is to perform a set of permutation tests (Chetty et al., 2009; Buchmueller et al., 2011; Baron et al., 2020). I estimate equations 2-4 with a placebo independent variable. This is done for every quarter-year combination from 2008 to 2019. This approach results in 48 placebo estimates (12 years x 4 quarters). The distribution of the 48 placebo estimates and one actual estimate from equations 2-4 represents the sampling distribution of $\hat{\beta}_1$, $\hat{\alpha}_1$, $\hat{\delta}_1$.

Figure 5 shows the cumulative distribution function of the placebo and actual estimates on *covid, schclo,* and *sah,* respectively for the referral rate and screened-in report rate. The actual estimate on *covid* is not statistically different from all of the placebo estimates, whereas the actual effect of the school closures and stay-at-home order are statistically different from the placebo estimates. These permutation tests indicate that the estimated impacts from school closures and the stay-at-home order are unlikely to be a result of chance.

Figure 5: Cumulative Distribution Function of Estimates from Permutation Tests



Covid Economics 82, 23 June 2021: 10-48



Notes: These figures plot the cumulative distribution function of the beta coefficient for the 48 placebo and one actual estimate. The top graph plots the coefficients on *covid* from equation 2, the middle graph plots the coefficients on *schclo* from equation 3, and the bottom graph plots the coefficient on *sah* from equation 4. The left-hand graphs plot the coefficient for the total referral rate (per 1,000 children) and the right-hand graphs plot the coefficient for the screened-in report rate.

Covid Economics 82, 23 June 2021: 10-48

6. Conclusion

There are three major findings of this paper. First, in Colorado, the COVID-19 pandemic and subsequent policy responses resulted in a 15 percent decline in reporting in 2020, compared to 2019. The biggest decline occurred between April and June as a result of the stay-at-home order and initial shift to virtual schooling; however, delayed school openings and continued spreading of the Coronavirus kept reporting below pre-pandemic averages for the remainder of 2020. Using a model that accounts for economic hardships and harmful coping strategies brought about by the pandemic, an estimated 38,800 referrals went unreported and 2,200 victims went unnoticed. Applying these numbers to the whole country while taking into account Colorado's unusually high reporting rate, an estimated 1.38 million referrals and 112,000 victims may have gone unreported across the country during 2020.[19] These estimates can (and should) be verified in two to three years when NCANDS releases the national-level data.[20]

There are three reasons to verify these numbers after the pandemic ends. First, for documentation purposes, it is important to correctly quantify the detrimental impacts of the pandemic. Second, these numbers can be used to predict the extent of underreporting in the event of another national emergency that alters maltreatment and reporting, simultaneously. Finally, child abuse and neglect has lasting consequences on educational attainment, employment and earnings, and health (Slade & Wissow, 2007; Irigaray et al., 2013; Kalmakis & Chandler, 2015; Doyle & Aizer, 2018), so states should be making efforts to follow-up with the children who may have been missed. Having an accurate count for how many children may have experienced abuse or neglect during the pandemic allows states to know when their follow-up efforts have reached all potential candidates.

The second key finding is that the stay-at-home order, school closings, and COVID-19 national emergency all substantially reduced child maltreatment referrals and screened-in reports. Many child welfare experts and governors understand they are missing the opportunity to protect children (NGA, 2020); however, some experts disagree and claim that the pandemic is filtering out the

[19] First the 38,800 unreported referrals are multiplied by 51. Next, 1,978,800 is divided by 1.43 because Colorado's report rate is 1.43 times higher than the typical state. The number of victims is not rescaled because Colorado's victim rate is similar to the average state (9.7 victims per 1,000 children versus 8.9 victims per 1,000 children).

[20] If one were to scale early predictions from Baron et al., 2020 and Rapaport et al., 2020 to the full year, they would estimate that 833,000 to 1.06 million referrals went unreported.

flood of unsubstantiated reports.[21] In Colorado, I do not find evidence that the composition of substantiated and unsubstantiated reports changes. Moreover, even if the pandemic filtered out unsubstantiated reports, it is unclear whether this is a good thing for child welfare. Ultimately, the answer depends on the state and what supports are provided to children and families of unsubstantiated reports. For example, in Colorado, children and families of unsubstantiated reports can be referred to other services, so a call to the hotline may connect families with needed resources (CDHS, nd b). In this case, fewer calls, regardless of the disposition, is concerning.

While child welfare experts predicted that more severe child maltreatment would arise from the pandemic (Hofmann, 2021), and some research has found that the proportion of ER visits from child abuse and neglect almost doubled, relative to 2019 proportions (Swedo et al., 2020), in Colorado, I do not find evidence of this hypothesis yet. Whether this is a limitation of the data or a glimmer of hope is unclear. Identifying victims from hotline calls is especially difficult in a state that screens out the majority of referrals and substantiates so few reports. Stephens-Davidowitz (2013) proposed two clever ways to try to identify victims of maltreatment. One method relies on using fatality counts (i.e. extreme cases of child maltreatment) and the other relies on using Google searches including terms like "child abuse and neglect." Unfortunately, this study is not able to employ either method. First, CDHS does not provide real-time data on fatalities. Second, Google searches related to child abuse and neglect in 2020 saw a substantial uptick the first week of March, prior to the pandemic. This uptick follows the release of the Netflix true crime miniseries documentary, "The Trials of Gabriel Fernandez," which was released February 26th. This limitation is an area that future research should continue to address as different types of maltreatment victims require different interventions.[22]

Finally, I find that the referral rate for neglect decreased by less in counties that were more likely to comply with the stay-at-home order relative to counties that were less likely to comply.[23] This result provides suggestive evidence of parental burnout and is supported by early research in the medical literature. For example, childhood injuries occurring in the home or on bicycles increased relative to sport and playground injuries during the pandemic (Bram et al., 2020). In

[21] For an example, see the opinion piece, "National Opinion: COVID-19 is not leading to more child abuse, it's cleaning the 'pollution' of false reports," published in the Arizona Daily Star on September 4, 2020.
[22] Fatality data during 2020 will be available through NCANDS for all states in two to three years.
[23] Bullinger et al. (2020) come to a similar conclusion in Indiana.

Covid Economics 82, 23 June 2021: 10-48

addition to targeting support and services to communities with historical records of substantiated cases, agencies need to target support and services to communities hit hard by the pandemic.

Eventually the pandemic will be a thing of the past; however, the findings of this study have implications beyond the pandemic. These findings quantify another hardship brought about by the pandemic: underreporting child maltreatment. The prevalence of underreporting highlights the role mandatory reporters play in detecting child maltreatment. These results can be used to inform policy decisions related to underreporting, mandatory reporting, and training. For example, states might want to consider additional ways to detect child maltreatment that do not rely on mandatory reporters to prepare for future events that may limit interactions between children and mandatory reporters. These results also speak to resource allocation for intervention after a pandemic. Based on findings from this paper, states should target resources to assist neglected children. For example, states may want to allocate additional funding to address the consequences of neglect. This paper can also inform policy decisions related to future pandemic responses. While school closures and stay-at-home orders reduced the spread of Coronavirus (Auger et al., 2020; Castillo et al., 2020), policymakers also have to consider the impact such policies had on child maltreatment reporting to design even better responses in the future. For example, the Department of Education in Maine provided an updated guide for teachers and others who care for children to detect maltreatment virtually (Maine DOE, 2020). Finally, this paper can be used as a reference to understand how events that alter maltreatment and reporting simultaneously impact child maltreatment referrals and substantiation rates in the data. Ultimately, fluctuations in the data seem to be more reflective of reporting than actual incidences.

References

ACF - U.S. Department of Health & Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. (2021). *Child Maltreatment 2019.* Available at https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/child-maltreatment.

Auger K.A., Shah S.S., Richardson T., et al. (2020). Association Between Statewide School Closure and COVID-19 Incidence and Mortality in the US. *JAMA,* 324(9):859–870. doi:10.1001/jama.2020.14348

Baron, E., Goldstein, E., & Wallace, C. (2020). Suffering in Silence: How COVID-19 school closures inhibit reporting of child maltreatment. *Journal of Public Economics, 190,* 104258.

Bram, J. T., Johnson, M. A., Magee, L. C., Mehta, N. N., Fazal, F. Z., Baldwin, K. D., Riley, J., & Shah, A. S. (2020). Where Have All the Fractures Gone? The Epidemiology of Pediatric Fractures During the COVID-19 Pandemic. *Journal of Pediatric Orthopedics*, 40(8): 373-379. https://doi.org/10.1097/BPO.0000000000001600

Brodeur, A., Clark, A.E., Fleche, S., & Powdthavee, N. (2021). COVID-19, lockdowns and well-being: Evidence from Google Trends. *Journal of Public Economics*, *193,* 104346. https://doi.org/10.1016/j.jpubeco.2020.104346

Brown, D., & De Cao, E. (2020). *Child Maltreatment, Unemployment, and Safety Nets* (SSRN Scholarly Paper ID 3543987). Social Science Research Network. https://doi.org/10.2139/ssrn.3543987

Buchmueller, T.C., DiNardo, J., Valletta, R.G. (2011). The effect of an employer health insurance mandate on health insurance coverage and the demand for labor: evidence from Hawaii. *American Economic Journal of Economic Policy*, 3 (4): 25–51.

Bullinger, L., Raissian, K., Feely, M., & Schneider, W. (2020). The neglected ones: time at home during COVID-19 and child maltreatment. *SSRN Working Paper*. http://dx.doi.org/10.2139/ssrn.3674064

Cabrera-Hernández, Francisco and Padilla-Romo, María. (2020). Hidden Violence: How COVID-19 School Closures Reduced the Reporting of Child Maltreatment, No 2020-02, Working Papers, University of Tennessee, Department of Economics, https://EconPapers.repec.org/RePEc:ten:wpaper:2020-02

Callahan, K. and Mink, C. (2020, May 7). Child abuse hotline calls are down during COVID-19, but abuse fears are up. *Center for Health Journalism.* https://centerforhealthjournalism.org/2020/05/05/child-abuse-hotline-calls-are-down-during-covid-19-abuse-fears-are.

Cameron, A.C., Gelbach, J.B., Miller, D.L. (2008). Bootstrap-based improvements for inference with clustered errors. *Rev. Econ. Stat.*, 90 (3): 414–427.

Castillo, R., Staguhn, E., Weston-Farber, E. (2020). The effect of state-level stay-at-home orders on COVID-19 infection rates. *American Journal of Infection Control,* 48(8): 958-960. https://doi.org/10.1016/j.ajic.2020.05.017

CDHS. (2020). Colorado Child Abuse and Neglect Hotline 2020 Data Shows 13 Percent Drop in Calls. https://www.prnewswire.com/news-releases/colorado-child-abuse-and-neglect-hotline-2020-data-shows-13-percent-drop-in-calls-301220637.html

CDHS (nd)a. Community Performance Center. http://www.cdhsdatamatters.org/data-by-topic.html

CDHS (nd)b. What happens after you call? https://www.co4kids.org/what-happens-after-you-call

CDPHE (nd). Outbreak Data. https://covid19.colorado.gov/covid19-outbreak-data. Accessed March 24, 2021.

Chetty, R., Looney, A., Kroft, K. (2009). Salience and taxation: theory and evidence. *Am. Econ. Rev.*, 99 (4):1145–1177.

Doyle, J. & Aizer, A. (2018). Economics of child protection: maltreatment, foster care, and intimate partner violence. *Annual Review of Economics*, 10: 87-108.

Ferman, B., Pinto, C., (2019). Inference in differences-in-differences with few treated groups and heteroskedasticity. *Rev. Econ. Stat,* 101 (3): 452–467.

Fitzpatrick, M., Benson, C., & Bondurant, S. 2020. Beyond Reading, Writing, and Arithmetic: The Role of Teachers and Schools in Reporting Child Maltreatment. *NBER Working Paper*.

Grewe, L. (2020). Stay-at-home: what you can and cannot do. *KKTV11 News.* https://www.kktv.com/content/news/STAY-AT-HOME-What-you-can-and-cannot-do-569122861.html

Griffith, A. (2020). Parental Burnout and Child Maltreatment During the COVID-19 Pandemic. *Journal of Family Violence*. https://doi.org/10.1007/s10896-020-00172-2

Hofmann, M. (2021). Child welfare officials see increase in reports of serious abuse. *Herald Standard.*

Irigaray, Tatiana Quarti, Pacheco, Janaína Barbosa, Grassi-Oliveira, Rodrigo, Fonseca, Rochele Paz, Leite, José Carlos de Carvalho, & Kristensen, Christian Haag. (2013). Child maltreatment and later cognitive functioning: a systematic review. *Psicologia: Reflexão e Crítica*, 26(2): 376-387.

Kalmakis, Karen and Genevieve Chandler. (2015). Health consequences of adverse childhood experiences: a systematic review. *American Association of Nurse Practitioners*, 27: 457-465.

Leslie, E. & Wilson, R. (2020). Sheltering in place and domestic violence: Evidence from calls for service during COVID-19. *Journal of Public Economics, 189,* 104241. https://doi.org/10.1016/j.jpubeco.2020.104241

MacFarlane, S., Yarborough, R., & Jones, S. (2020). Child advocates concerned neglect, abuse might be going unreported during pandemic. *NBC4 Washington.*

MacKinnon, J.G., Webb, M.D. (2017). Wild bootstrap inference for wildly different cluster sizes. *J. Appl. Econ*, 32 (2): 233–254.

MacKinnon, J.G., Webb, M.D. (2018). The wild bootstrap for few (treated) clusters. *Econ. J.*, 21 (2): 114–135.

Maine DOE. (2020). Priority Notice: Spotting signs of child abuse and neglect during the COVID-19 emergency: An updated guide for educational professionals and others who care for Maine children. Available at https://mainedoenews.net/2020/04/15/priority-notice-spotting-signs-of-child-abuse-and-neglect-during-the-covid-19-emergency-an-updated-guide-for-educational-professionals-and-others-who-care-for-maine-children/

Mikolajczak, M., Gross, J. J., & Roskam, I. (2019). Parental Burnout: What Is It, and Why Does It Matter? *Clinical Psychological Science*, *7*(6), 1319-1329. https://doi.org/10.1177/2167702619858430

NGA - National Governor's Association. (2020). Addressing the decline in child abuse reports and supporting child well-being.

Ortiz R, Kishton R, Sinko L, et al. (2021). Assessing Child Abuse Hotline Inquiries in the Wake of COVID-19: Answering the Call. *JAMA Pediatr.* doi:10.1001/jamapediatrics.2021.0525

Pollard M.S., Tucker J.S., Green H.D. (2020). Changes in Adult Alcohol Use and Consequences During the COVID-19 Pandemic in the US. *JAMA Network Open.*3(9):e2022942. doi:10.1001/jamanetworkopen.2020.22942

Quander, M. (2020, September 2). More than 60% drop in calls to child abuse hotline spark safety concerns. *Newbreak.* https://www.newsbreak.com/news/2052707885364/more-than-60-drop-in-calls-to-child-abuse-hotline-spark-safety-concerns.

Rapoport, E., Reisert, H., Schoeman, E, & Adesman, A. 2020. Reporting of child maltreatment during the SARS-CoV-2 pandemic in New York City from March to May 2020. *Child Abuse & Neglect,* 104719. https://doi.org/10.1016/j.chiabu.2020.104719

Schmidt, S. & Natanson, H. (2020). With kids stuck at home, ER doctors see more severe cases of child abuse. *The Washington Post.*

Slade, E. P., & Wissow, L. S. (2007). The influence of childhood maltreatment on adolescents' academic performance. *Economics of education review*, 26(5), 604–614.

Stephens-Davidowitz, S. (2013). Unreported victims of an economic downturn. Retrieved from Seth Stephens-Davidowitz website: https://static. squarespace. com/static/51d894bee4 b01caf88ccb4f3

Trading Economics. (nd). United States Unemployment Rate. https://tradingeconomics.com/united-states/unemployment-rate#:~:text=Unemployment%20Rate%20in%20the%20United,percent%20in%20May%20of%201953. Accessed March 24, 2021.

WCTV. (2020, May 25). Advocates express concerns about children falling through the cracks. https://chsfl.org/blog/advocates-express-concerns-about-children-falling-through-the-cracks/.

Weiner, D., Heaton, L., Stiehl, M., Chor, B., Kim, K., Heisler, K., Foltz, R., & Farrell, A. (2020). Chapin Hall issue brief: COVID-19 and child welfare: Using data to understand trends in maltreatment and response. Chicago, IL: Chapin Hall at the University of Chicago.

WHO – World Health Organization. (nd). Child maltreatment and alcohol abuse. https://www.who.int/violence_injury_prevention/violence/world_report/factsheets/fs_child.pdf.

Appendix

Appendix Table 1: State-Level Differences in Child Maltreatment Reporting and Economic Conditions between 2019 and 2020

Year	2019					2020				
Quarter	Jan - Mar	Apr - Jun	Jul - Sep	Oct - Dec	TOTAL	Jan - Mar	Apr - Jun	Jul - Sep	Oct - Dec	TOTAL
Child Maltreatment Variables										
Total Referrals Received	28626	28281	28549	29722	115178	29819	19577	24482	24280	98158
Screened-in	9896	9782	9850	9422	38950	9727	7190	8856	8348	34121
Screened-out	18730	18499	18699	20300	76228	20092	12387	15626	15932	64037
Total Allegations of Maltreatment	15333	15195	15688	14875	61091	15457	12020	14757	13019	55253
Substantiated	3487	3397	3649	3197	13730	3426	3014	3322	2922	12684
Unsubstantiated	11846	11798	12039	11678	47361	12031	9006	11435	10097	42569
Neglect Allegations	10865	10785	11349	10577	43576	10865	8955	10660	9386	39866
Physical Abuse Allegations	2623	2528	2548	2694	10393	2765	1693	2363	2045	8866
Sexual Abuse Allegations	1106	1122	1106	938	4272	1105	842	1091	935	3973
Substantiated Neglect	2786	2661	2940	2541	10928	2718	2462	2662	2335	10177
Substantiated Physical Abuse	289	336	310	295	1230	336	249	306	254	1145
Substantiated Sexual Abuse	311	299	310	270	1190	268	221	269	238	996
Economic Conditions										
Unemployment rate	3.10	2.80	2.63	2.50		3.40	11.00	6.83	7.07	
Employment-population ratio	66.93	67.13	67.57	67.70		66.67	59.87	62.37	63.50	
Alcohol Purchased (gallons per capita)					2.78					2.99
Percent Change between 2020 and 2019										
Child Maltreatment Variables										
Total Referrals Received	0.04	-0.31	-0.14	-0.18	-0.15					
Screened-in	-0.02	-0.26	-0.10	-0.11	-0.12					
Screened-out	0.07	-0.33	-0.16	-0.22	-0.16					
Total Allegations of Maltreatment	0.01	-0.21	-0.06	-0.12	-0.10					

CEPR PRESS

Covid Economics 82, 23 June 2021: 10-48

Substantiated	-0.02	-0.11	-0.09	-0.09	-0.08
Unsubstantiated	0.02	-0.24	-0.05	-0.14	-0.10
Neglect Allegations	0.00	-0.17	-0.06	-0.11	-0.09
Physical Abuse Allegations	0.05	-0.33	-0.07	-0.24	-0.15
Sexual Abuse Allegations	0.00	-0.25	-0.01	0.00	-0.07
Substantiated Neglect	-0.02	-0.07	-0.09	-0.08	-0.07
Substantiated Physical Abuse	0.16	-0.26	-0.01	-0.14	-0.07
Substantiated Sexual Abuse	-0.14	-0.26	-0.13	-0.12	-0.16
Economic Conditions					
Unemployment rate	0.10	2.93	1.59	1.83	
Employment-population ratio	0.00	-0.11	-0.08	-0.06	
Alcohol Purchased (gallons per capita)					0.08

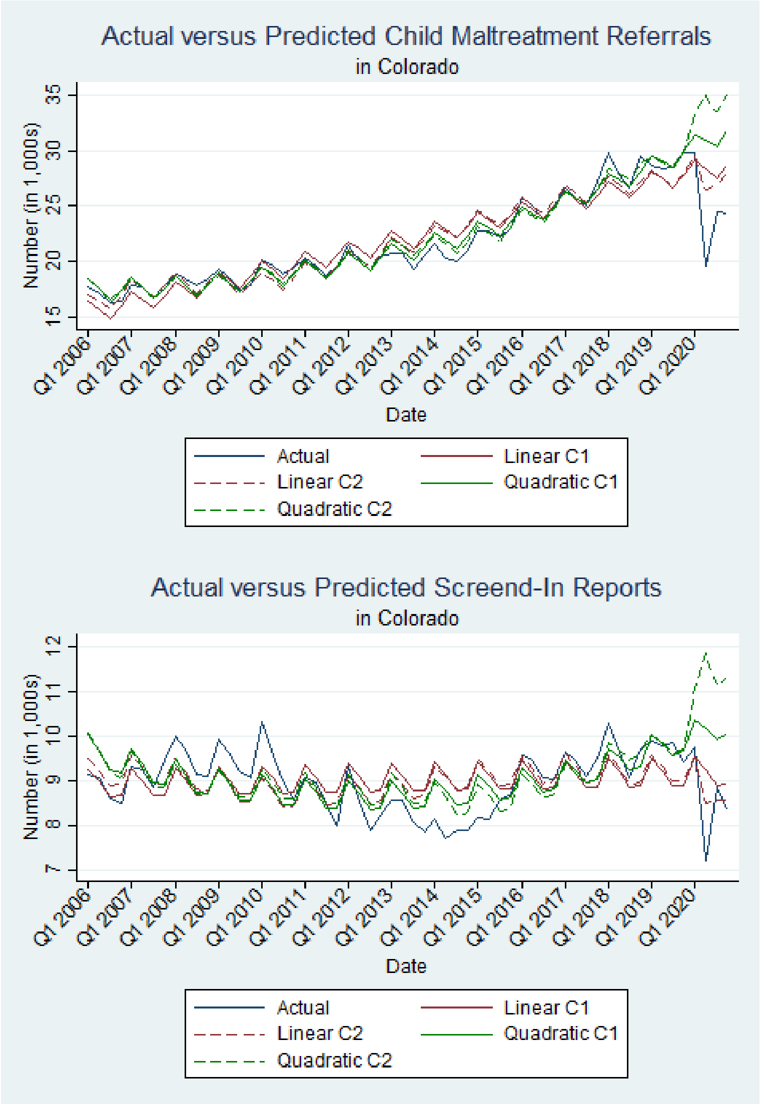
Notes: This table reports the number of referrals, screened-in and screened-out reports, allegations (including disposition), and allegations by maltreatment type for the state of Colorado for each quarter and the full year in 2019 and 2020. In addition, some economic conditions (unemployment rate and employment-population ratio) and a measure of a coping technique (alcohol sales) is included. The bottom panel reports the percent change between 2020 and 2019 for quarter and the year for each variable.

Appendix Table 2: Summary Statistics of Child Maltreatment Reporting and Economic Conditions in Colorado

	Mean (N=3,328)	Std. Dev.	2019 Average (N=256)	2020 Average (N=256)	p-value
Child Maltreatment Variables (per 1,000 children)					
Total Referrals	18.03	9.66	24.66	20.16	0.00
Screened-in	7.97	5.00	7.83	7.03	0.05
Screened-out	10.06	7.21	16.83	13.14	0.00
Total Allegations of Maltreatment	12.81	9.23	12.53	11.70	0.27
Substantiated	3.12	3.46	2.61	2.57	0.86
Unsubstantiated	9.69	7.34	9.92	9.14	0.21
Neglect Allegations	8.45	6.73	8.72	7.80	0.09
Physical Abuse Allegations	2.37	2.04	2.20	2.08	0.48
Sexual Abuse Allegations	0.90	1.12	0.79	0.87	0.33
Substantiated Neglect	2.26	2.73	1.97	1.83	0.51
Substantiated Physical Abuse	0.36	0.67	0.25	0.25	0.85
Substantiated Sexual Abuse	0.23	0.58	0.19	0.26	0.19
Economic Conditions					
Unemployment Rate	5.46	2.82	2.86	6.41	0.00
Employment to Population Ratio	50.55	9.67	53.68	48.95	0.00
Child Population (0 to 17)	19471	39899	19680	19665	1.00

Notes: This table provides summary statistics for child maltreatment reporting and economic conditions across all counties in Colorado from 2008 to 2020. The mean and standard deviation are given for all 3,328 observations (13 years x 4 quarters x 64 counties) in columns 1 and 2. The 2019 and 2020 averages and corresponding p-value from a t-test are provided in columns 3-5. All averages for the child maltreatment variables are provided as rates, per 1,000 children, so the average of 18.03 means that in a typical quarter a county received 18.03 referrals, per 1,000 children.

Appendix Figure 1: Actual versus Predicted Child Maltreatment Reporting using Linear and Quadratic Polynomial



Notes: This figure plots the counterfactual estimates from equation 1 using a linear and quadratic polynomial, as opposed to the preferred cubic specification. The linear model is less predictive than the quadratic, and the quadratic model is similar to the cubic model.

Appendix Table 3: Summary Statistics of Child Maltreatment Reporting by COVID-19 Cases

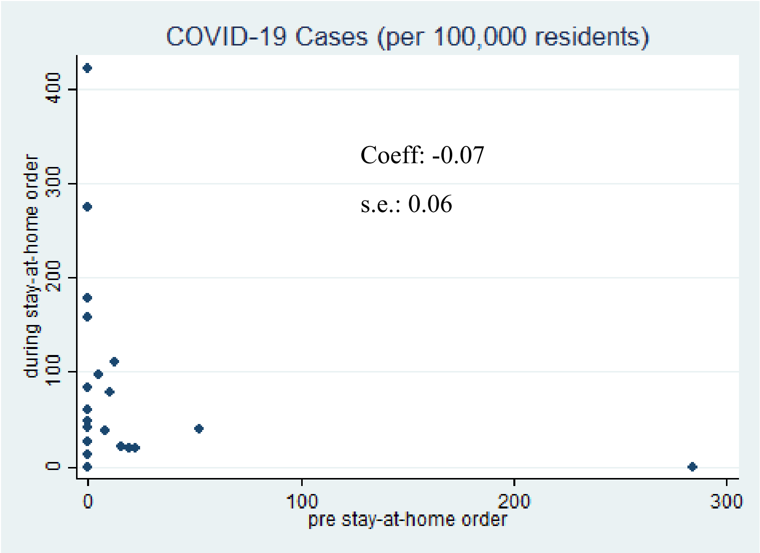
	2019			2020			
	(1) Non-COVID-19 Counties	(2) COVID-19 Counties	p-value	(3) Non-COVID-19 Counties	(4) COVID-19 Counties	p-value	Unconditional DD Estimate (4-2-(3-1))
Child Maltreatment Variables (per 1,000 children)							
Total Referrals	24.83	23.62	0.17	20.23	19.77	0.61	0.76
Screened-in	7.81	7.96	0.71	7.03	6.99	0.90	-0.20
Substantiated Reports	2.66	2.34	0.31	2.62	2.24	0.25	-0.06
Neglect Allegations	8.73	8.62	0.84	7.76	8.05	0.57	0.40
Physical Abuse Allegations	2.22	2.06	0.36	2.13	1.72	0.02	-0.26
Sexual Abuse Allegations	0.78	0.80	0.79	0.89	0.78	0.26	-0.13
Economic Conditions							
Unemployment Rate	2.89	2.64	0.02	6.33	6.93	0.24	0.85
Employment to Population Ratio	53.75	53.24	0.62	48.79	49.97	0.26	1.70
Child Population	6009	103223	0.00	5988	103247	0.00	45.36
COVID-19 Cases (per 100,000 residents)							
pre stay-at-home order				0.00	47.61	0.15	
during stay-at-home order				23.75	47.04	0.17	
post stay-at-home order				59.23	26.84	0.31	
Number of Counties	55	9		55	9		

Notes: This table reports the average child maltreatment referral, screened-in, and substantiated report rate, and economic conditions across all counties and quarters in 2019 and 2020 based on the counties' COVID-19 caseload. Nine counties ("treated") reported COVID-19 cases prior to the stay-at-home order, whereas 55 counties ("control") reported no cases prior to the stay-at-home order. The p-value for t-tests between the COVID-19 and non-COVID-19 counties are reported for each year, and a crude DD is provided.

Covid Economics 82, 23 June 2021: 10-48

CEPR PRESS

Appendix Figure 2: Relationship between COVID-19 Cases Before and During the Stay-at-home Order



Notes: This figure plots the number of COVID-19 cases (per 100,000 residents) before the stay-at-home order against the number of cases during the month of the stay-at-home order. The negative regression adjusted coefficient (-0.07) suggests places with higher pre-stay-at-home order COVID-19 cases might have been more compliant to the stay-at-home order.

Appendix Table 4: Robustness Analyses for Screened-in Reports

		(1) Main Results	(2) Post 2010	(3) Exclude Denver Metro-area	(4) Include Economic Controls
Panel A: Rate (per 1,000 children)	COVID-19	-0.079	-0.079	-0.024	0.118
		(0.710)	(0.708)	(0.801)	(0.795)
	School Closure	-1.838**	-1.814**	-1.711*	-1.883*
		(0.846)	(0.831)	(0.966)	(0.961)
	Stay-at-home	-3.273**	-3.291**	-3.146*	-3.514**
		(1.479)	(1.466)	(1.681)	(1.759)
Panel B: Log	COVID-19	-0.004	0.003	0.012	0.031
		(0.095)	(0.094)	(0.106)	(0.104)
	School Closure	-0.201*	-0.195*	-0.152	-0.176
		(0.121)	(0.117)	(0.137)	(0.134)
	Stay-at-home	-0.382*	-0.375*	-0.309	-0.347
		(0.211)	(0.205)	(0.238)	(0.242)
Panel C: Levels	COVID-19	-19.531*	-21.527*	-8.392	-15.955
		(11.504)	(11.230)	(9.196)	(11.007)
	School Closure	-45.600***	-43.433***	-21.899**	-41.476**
		(16.603)	(15.174)	(10.072)	(16.030)
	Stay-at-home	-89.454***	-86.960***	-44.373**	-84.199***
		(29.659)	(27.306)	(18.952)	(28.668)
	Observations	3,328	2,560	2,860	3,328

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors, clustered at the county-by-quarter level, in parentheses. Each column indicates a separate regression analysis. The first column provides the main results. The second column is post-2010. The third column excludes the Denver metro-area, and the fourth column includes controls for economic conditions (i.e. the unemployment rate and employment-to-population ratio). Each row represents a separate regression analysis for the three independent variables of interest, so row 1 reports the coefficient from equation 2 where *covid* is the independent variable of interest. Row 2 reports the coefficient from equation 3 where *schclo* is the independent variable of interest, and row 3 reports the coefficient from equation 4 where *sah* is the independent variable of interest. There are 3 separate panels as well. Panel one reports the effect on the screened-in report rate, per 1,000 children. The second panel reports the effect on the log screened-in reports, and the third panel reports the effect on the total screened-in reports (level). Each regression includes year, county, and quarter fixed effects.

Covid Economics 82, 23 June 2021: 10-48

Appendix Table 5: Robustness Analyses for Substantiated Reports

		(1) Main Results	(2) Post 2010	(3) Exclude Denver Metro-area	(4) Include Economic Controls
Panel A: Rate (per 1,000 children)	COVID-19	-0.876	-0.947	-1.077	-0.855
		(0.715)	(0.701)	(0.817)	(0.749)
	School Closure	0.615	0.412	0.870	0.933
		(0.966)	(0.961)	(1.106)	(1.150)
	Stay-at-home	0.628	0.227	0.939	1.236
		(1.764)	(1.748)	(2.028)	(2.166)
Panel B: Log	COVID-19	-0.046	-0.063	-0.079	-0.016
		(0.154)	(0.146)	(0.171)	(0.161)
	School Closure	0.206	0.179	0.281	0.300
		(0.201)	(0.191)	(0.225)	(0.227)
	Stay-at-home	0.315	0.261	0.428	0.516
		(0.356)	(0.338)	(0.401)	(0.416)
Panel C: Levels	COVID-19	-5.794	-7.053	-5.118	-5.196
		(8.598)	(7.548)	(7.281)	(8.558)
	School Closure	-5.966	-5.429	0.329	-3.918
		(8.473)	(6.771)	(6.232)	(8.431)
	Stay-at-home	-13.249	-13.255	-1.623	-9.381
		(15.645)	(12.636)	(12.047)	(15.559)
	Observations	3,328	2,560	2,860	3,328

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors, clustered at the county-by-quarter level, in parentheses. Each column indicates a separate regression analysis. The first column provides the main results. The second column is post-2010. The third column excludes the Denver metro-area, and the fourth column includes controls for economic conditions (i.e. the unemployment rate and employment-to-population ratio). Each row represents a separate regression analysis for the three independent variables of interest, so row 1 reports the coefficient from equation 2 where *covid* is the independent variable of interest. Row 2 reports the coefficient from equation 3 where *schclo* is the independent variable of interest, and row 3 reports the coefficient from equation 4 where *sah* is the independent variable of interest. There are 3 separate panels as well. Panel one reports the effect on the substantiated report rate, per 1,000 children. The second panel reports the effect on the log substantiated reports, and the third panel reports the effect on the total substantiated reports (level). Each regression includes year, county, and quarter fixed effects.

Covid Economics Issue 82, 23 June 2021

Covid-19 impact on artistic income[1]

Alexander Cuntz[2] and Matthias Sahli[3]

Date submitted: 10 June 2021; Date accepted: 15 June 2021

This paper assesses the impact of the pandemic crisis on self-employed income among artists resident in Germany. Using unique data from the latest available public insurance records, we show that musicians and performing artists are among the most vulnerable groups, and that writers, on average, are relatively less impacted. Moreover, the paper looks at the impact of the 2020 crisis on income differences by gender, career stages and regions, and it investigates the effect of specific non-pharmaceutical, public intervention implemented in German states.

Covid Economics 82, 23 June 2021: 49-69

1 The authors gratefully acknowledge data donation by the Kuenstlersozialkasse. They also would like to thank Carsten Fink, Andreas Kissling and seminar participants at the Onassis Foundation, Athens, for valuable advice and comments on earlier versions of the paper. The views expressed are those of the authors, and do not necessarily reflect he views of the World Intellectual Property Organization or its member states.

2 Head of the Creative Economy Section, World Intellectual Property Organization, Department for Economics and Data Analytics.

3 Research Fellow & PhD Candidate, World Intellectual Property Organization, Department for Economics and Data Analytics & University of Neuchâtel.

Covid Economics 82, 23 June 2021: 49-69

1 Motivation

Data and timely evidence for well-informed policy in the cultural and creative sectors are hard to come by in a state of emergency such as the current global pandemic. In a time of crisis, policy makers might not be able to build on previous experiences and learn from historic evidence. Still, there is a need to allocate public resources and support as well as identify and reach out to most vulnerable groups among artists. Based on the latest available data released by official sources in March 2021, this paper makes an attempt to assess the Covid-19 impact on self-employed income from artistic practice and among artists located in Germany.

We find that the pandemic crisis impacts artists in creative and cultural sectors differently, net income losses ranging between 2 to 13 percent. More precisely, our results indicate that musicians and performing artists are among the most vulnerable groups in terms of income losses in 2020, and that some losses may depend on the specific non-pharmaceutical, public intervention implemented in German states. Furthermore, we can show that gender income differences and differences at different career stages largely prevail over the crisis and predate the 2020 outbreak, and that artists in rural areas are no less affected than those in urban areas.

We contribute to a growing number of economic studies assessing the pandemic's impact and the impact of specific containment measures (Baldwin and Di Mauro, 2020; Cusmano and Raes, 2020; von Bismarck-Osten et al., 2020), in particular in the Arts and Culture (Buchholz et al., 2020; Jacobs et al., 2020). For example, topical research by the PEC centre documents a contraction of the UK labor market and fewer hours

Covid Economics 82, 23 June 2021: 49-69

worked in these sectors.[1] Using data from the national Labour Force Survey (LFS), the study estimates a loss of 55,000 jobs which equals a 30 percent decline in music, performing and visual arts between the first and the third quarter of 2020. In addition, a number of recent studies by the European Parliament show similar contractions of these sectors and beyond across the EU and based on Eurostat data (De Vet et al., 2021).[2] Moreover, EU studies track and monitor national public support measures for creators and identify the many non-standard workers such as self-employed and part-time workers as the most vulnerable group in these sectors during the first wave of the 2020 pandemic. Notably, our research is backed by the latest available income data reported to an official public insurance scheme for artists located in Germany and it covers all of 2020. Different to many other Covid-19 impact studies, our research does not rely on income forecasts based on historic data, and it does not suffer potential bias from survey responses and the adequacy of sampling techniques. Moreover, we trust that more concise estimates of income losses experienced by artists in 2020 may enable policymakers to better target public support in 2021, and help tailor financial and other support for most vulnerable groups and regions.

The paper structures as follows. Section 2 describes the unique dataset and the limitations of the study, section 3 sets up a simple empirical framework, and section 4 presents main findings from the analysis. Section 5 concludes.

2 Data and Study Limitations

Income data for around 190,000 artists comes from social insurance records of the 'Kuenstlersozialversicherungskasse' (KSK) and has been used in previous research on

[1] See this link to their blog.
[2] See the reports released here and there.

the financial health of creators (Cuntz, 2018; Kretschmer, 2005). The dedicated low-cost insurance scheme targets artists resident in Germany. It requires them to report self-employed income on an annual basis which they later also report to tax authorities. Applicants to the scheme self-identify as (self-employed) artists, and so, from a methodological perspective, there is no need for us to survey artists nor define sample criteria ex ante. In addition, once artists have opted-in, they self-select into one out of four artistic categories, i.e. fine arts, performing arts, music and writing/literature.

The aggregate KSK records we can access report average income/net revenue (mean) from artistic self-employment and the number of insured artists per group by gender, artistic category, age group and geography (NUTS-1 'Laender' or states). The data is available for four consecutive years of reported income 2017 to 2020 and thus it only accounts for the impact on income from first and second pandemic waves in 2020. Furthermore, the data distinguishes artists at early career stages from the total insured population. The KSK defines 'early stage' careers as the first three years of reporting artistic self-employment to them. Tentative analysis of the number of insured artists suggests that there is no substantial crisis impact as the total stock of insured persons does not change in 2020 over previous years (results not shown).

We complement the income data with regional information on non-pharmaceutical interventions (NPI) implemented in 2020 in order to contain the spread of COVID 19 during the first and second pandemic waves (Cheng et al., 2020), i.e. December 31, 2020. NPIs restricted movement, public gatherings, international travel as well as led to the closure of education institutions and retail stores in the 16 German Laender and at different points in time, introducing some limited spatio-temporal variations in the introduction of policy measures (Aravindakshan et al., 2020). The introduction of

Covid Economics 82, 23 June 2021: 49-69

NPIs, in turn, is linked to the spread of the pandemic over time in each of the Laender (Figure 1).

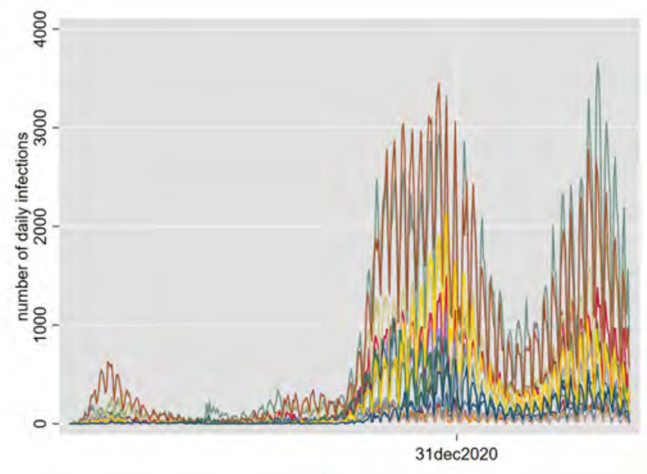


Figure 1: COVID-19 spread by region

Note: This figure shows the time trend in new daily infections for each of the 16 German Laender (NUTS-1), and for the first, second and third pandemic waves. The vertical line indicates the cut-off date for the reference period of the 2020 income data in the midst of the second wave. Source: RKI pandemic data

Several caveats apply to the KSK data. It does not provide any information on alternative sources of income for insured artists. For example, they will often cross-subsidize their artistic income and, in this way, might compensate for some self-employed income losses due to the pandemic via working multiple jobs and taking up regular employment, within and outside the arts. Moreover, as some of the pre-pandemic 2019 royalties were only distributed in 2020, this might flatten and upward bias the impact we observe on reported income. So, arguably, pandemic 2020 royalty changes will only be reflected in the forthcoming 2021 data. In addition, reported income in 2020 might suffer from bias due to funding support artists received from private and public sources. For example,

Covid Economics 82, 23 June 2021: 49-69

Covid Economics 82, 23 June 2021: 49-69

taxable funding on federal government level aiming to compensate for the substantial changes in revenues by artists and other self-employed workers might upward bias income figures (e.g. the so-called "November-/Dezemberhilfen" funds).[3] However, this main source of public funding in Germany does not seem relevant here as the distribution of funds was delayed until late January 2021 and thus it should not impact 2020 reported income.[4] Moreover, most financial support was based on the reimbursement of lost fixed costs which meant many self-employed were not eligible for these schemes. If our data would nevertheless be biased by funding and support measure, we could still treat and consider results as lower bound estimates of the pandemic impact.

Finally, other contextual information might be equally important to consider as a determinant of income changes. For example, the 2020 pandemic might serve as a catalyst to digital change and digital literacy among artists, based on recent industry surveys.[5] Arguably, trends towards more online distribution and cultural consumption as well as, related to that, alternative income sources and reworked business models go largely unobserved with the given data and can also moderate the income changes we observe in the next section.

3 Empirical Approach

We are interested in understanding how income in the different creative sectors is affected by the Covid-19 crisis in Germany, accounting for the various NPIs on the state-level. In an ideal empirical setting, a (quasi) natural experiment would randomly assign treatments, i.e. different sets of NPIs, to a group of treated artists and not assign

[3]See, for example, this or that link (in German).
[4]See, for example, this source (in German).
[5]See, for example, this source (in German).

them to a control group of similar artists. Looking at both groups, we could isolate the causal effect of NPIs on income in each state and creative sector, controlling for other relevant factors associated with the spread of the Covid-19 pandemic. Instead, the empirical approach begins with a straightforward income estimation, and baseline results are obtained from estimating the following equation:

$$Inc_{s_{ct}} = \alpha + \delta * (Post_{st}) + \boldsymbol{X} * (Control_{s_{ct}}) + \mu_c + \epsilon_{s_{ct}}, \tag{1}$$

where Inc describes the average income of insured artists in state c, sector s and year t. Baseline regressions include state-fixed effects (μ) and standard errors are clustered at the state-level. Further control variables are summarized by $\boldsymbol{X}$ and capture artist group characteristics such as gender, age cohorts and career stages. We thus run separate regressions in each sector in order to isolate the $Post_{st}$ year effect, and then test the statistical significance of the difference between the 2019 and 2020 coefficients.

In a second step, we combine all sectors and run a single regression that identifies post-2019 income effects, based on the baseline specification outlined in equation (1):

$$Inc_{s_{ct}} = \alpha + \sum_t \beta^t year + \delta * (2020_t \times Sector_s) + \boldsymbol{X} * (Control_{s_{ct}}) + \mu_c + \epsilon_{i_t}, \tag{2}$$

and where $Sector_s \in \{$Writing, Fine Art, Music and Performing Art$\}$. Here, again our coefficient of interest is represented by δ, i.e. the interaction term of sectors and the 2020_t year treatment variable, capturing income-effects of the 2020 pandemic year. Baseline regressions include year-fixed effects (β), state-fixed effects (μ), and controls ($\boldsymbol{X}$) as well as standard errors clustered at the state-level.

Covid Economics 82, 23 June 2021: 49-69

CEPR PRESS

Furthermore, as we estimate factor-variables interactions with year dummies, an meaningful approach is to interpret and graph income gaps using the margins-command (Stata). Margins are statistics calculated based on predictions from the above model. Thus we calculate margins based on the LHS of equation (1) as $\frac{\Delta E[y|sector,year]}{\Delta year}$, allowing the intercept and marginal effect (slope) of $post2019$ outcomes to be different for each sector, see for instance (Karaca-Mandic et al., 2012; Perraillon, 2021).

Finally, we refine our empirical strategy and estimate the income effects for different types and qualities of NPIs impacting each sector. We calculate post-2019 income effects as described above, but include an approximation of NPI intervention period lengths for each state (for example, the number of lockdown days). This variation allows to account for the heterogeneity in state-level responses to the Covid-19 pandemic in Germany. Hence, we add to the baseline equation (1) and the coefficient of interest δ an interaction term NPI_c in state c. We distinguish NPI treatments in upper and lower percentiles of intervention periods (<p25 and >p75) and test whether the effect of specific NPIs and lengths on 2020 income differs across sectors, with $NPI_c = 0$ and $NPI_c = 1$. So, the interaction term $\delta * (Post_t \times NPI_c)$ identifies this effect and we again run separate models for each sector in order to report the coefficient of interest.

4 Findings and Discussion

4.1 Baseline Results: Sectoral Income Gap

Our baseline results are presented in table 1 and table 2. In table 1 we run the regression for each sector separately and samples are restricted to years 2019-2020. In table

Covid Economics 82, 23 June 2021: 49-69

Covid Economics 82, 23 June 2021: 49-69

2, we report post-2019 interaction income effects for all sectors and based on the larger sample for years 2017-2020. All models control for gender, age cohorts, sectors and career stage effects on income. Both tables show results for our preferred specification including state-level fixed-effects, standard errors clustered at the state-level and year fixed-effects.

The results in table 1 suggest an economically significant, negative income effect in 2020 across all creative sectors and over the previous year. The effect is statistically significant at the 1%-level (or lower) for all sectors except writers. Income drops in 2020 vary across sectors as model (1) to (4) illustrate. We estimate the strongest decline for the group of performing artists, with an average income loss of about 1'998 Euro in 2020, 1'485 Euro for musicians, 1'390 Euro for fine artists and 280 Euro for writers. This loss of income from artistic self-employment is, as we are able to illustrate in the following, substantial when compared to the reported income in 2019 (median) of writers (18'900 Euro), fine artists (15'300 Euro), musicians (12'300 Euro) and performing artists (14'400 Euro).

Negative 2020 income effects continue to hold in our second model specification. Table 2 reports results. In each column, we present overall 2017-2020 sectoral differences in income levels, and interaction terms $\{2020 \times Sector_s\}$ capture post-2019 income changes in each sector over the average income losses experienced by writers in the same period (base). Estimated losses stay economically significant and robust, and relative magnitudes of sectoral effects are confirmed. Notably, linear regression models in table 2 can explain roughly 30 percent of the variance in the outcome variable (1), and, under the log-transformed outcome in model (2), goodness-of-fit measures R^2 increase to 35 percent.

CEPR PRESS

Table 1: Income Creative Sectors

	(1) inc	(2) inc	(3) inc	(4) inc
Post (Writing)	-280.9			
	(-0.74)			
Post (Fine Art)		-1390.8**		
		(-3.99)		
Post (Music)			-1485.3***	
			(-4.62)	
Post (Performing Art)				-1998.8***
				(-6.55)
Female	-4876.7***	-6039.2***	-2585.7**	-6237.9***
	(-9.04)	(-7.00)	(-3.56)	(-9.24)
Early Career	-1746.8*	-1045.8	-3029.3***	-2222.0**
	(-2.50)	(-1.24)	(-10.30)	(-3.17)
Constant	24941.3***	22586.8***	25122.0***	23931.9***
	(9.72)	(8.80)	(25.77)	(9.09)
N	616	617	601	576
R^2	0.170	0.281	0.172	0.318
N Groups	17	17	17	17
FE	Laender	Laender	Laender	Laender
Cluster SE	Laender	Laender	Laender	Laender
Sample	2019-2020	2019-2020	2019-2020	2019-2020

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the baseline results for the dependent variable *income* in model (1)-(4) and separate estimations for the four sectors. All models include age cohort controls which are not reported. The sample is calculated based on reported self-employed income for 2019-2020. All estimates derive from ordinary-least-squares (OLS) models.

We visualize our main results in figure 2. Notably, writers such as self-employed and freelance journalists, authors or other publicists in this artistic group are exceptional with regard to their income losses. The results indicate that they were able to maintain or slightly increase their self-employed income in the 2020 pandemic year. Arguably, our point estimates for this sector indicate that writers more than other artists were able

Covid Economics 82, 23 June 2021: 49-69

Table 2: Income Creative Sectors (overall)

	(1) inc	(2) log(inc)
Writing (base)		
Fine Art	-3205.6***	-0.200***
	(-6.55)	(-6.15)
Music	-6670.0***	-0.416***
	(-15.10)	(-17.40)
Performing Art	-4294.1***	-0.275***
	(-8.70)	(-11.16)
2020 $\times FineArt$	-671.6	-0.0314
	(-1.35)	(-1.28)
2020 $\times Music$	-1199.6**	-0.106***
	(-3.43)	(-4.15)
2020 $\times PerformingArt$	-1376.1*	-0.0990**
	(-2.92)	(-3.75)
Female	-4486.8***	-0.264***
	(-14.44)	(-18.69)
Early Career	-2336.6***	-0.269***
	(-8.22)	(-14.67)
Constant	26814.9***	10.66***
	(21.81)	(138.10)
N	4816	4816
R^2	0.296	0.355
N Groups	17	17
Year FE	Yes	Yes
FE	Laender	Laender
Cluster SE	Laender	Laender
Sample	2017-2020	2017-2020

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the baseline results (overall) for the dependent variable *income* in model (1) and $log(income)$ in (2). All models include further age cohort and year-effects controls. The sample is calculated based on reported self-employed income for 2017-2020. All estimates are derived from ordinary-least-squares (OLS) models.

Covid Economics 82, 23 June 2021: 49-69

Covid Economics 82, 23 June 2021: 49-69

to continue their work online, in a time when theaters and concert venues were closing due to the restrictions imposed by NPIs.[6] Figure 2 also reveals sectoral differences pre-dating the crisis as average self-employed income is highest for writers over the entire period of observation, followed by fine artists, musicians and performing artists.

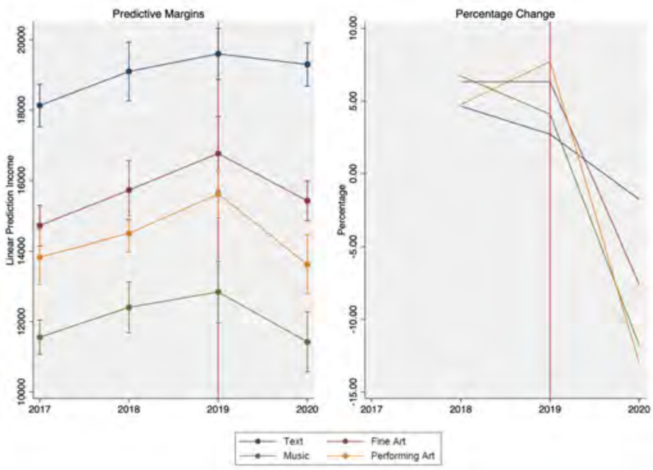


Figure 2: Income by Creative Sectors
Note: The left-hand panel shows absolute income levels by creative sectors, with writing (blue), fine arts (red), music (green) and performing arts (yellow). The panel on the right presents relative income changes from one period to the next.

4.2 Impact on Equal Pay, Career Stages and Regions

We can further inspect sectoral impacts of the Covid-19 pandemic along several important dimensions. For example, we can test if the pandemic had an impact on the existing gender income gap as well as test for the changes in income at different career stages and among artists resident urban or rural areas.

[6]NPIs are discussed in greater detail in section 4.3 below.

CEPR PRESS

First, as Figure 3 illustrates the impact of the pandemic does not uniformly affect women and men across all sectors.[7] Here, it is interesting to note that female writers seem to outperform trends among male counterparts in 2020. Put differently, women in the sector see income growth over their 2019 levels, while men are losing some of their income over the same period. This is a notable result as at odds with the notion developed in many other studies on Covid-19 impact (Xue and McMunn, 2021). Findings there suggest that women, more often than men, are taking over additional household hours during the pandemic, e.g. when home schooling kids etc. Arguably, this should also be reflected in hours worked in professional lives as well as the relative changes in income observed for each group. At the same time, it could be that female writers have more flexible work arrangements in the first place, for example, in terms of working on a less fixed time schedule or working from home/remotely in pre-pandemic times, shielding them from some of the losses. At large, however, diverging trends have little overall effect on the existing, pre-crisis pay gap (i.e. the average sectoral income gap, to be seen in table 1 coefficient $female$, is nonetheless significant and negative). In most sectors, women are equally losing self-employed income from artistic practice. Interestingly, in the performing arts, pandemic losses in 2020 should be assessed against the observation that women did not fully participate in the income growth from pre-pandemic years, mostly benefiting male performers in the same sector.

Second, Figure 4 depicts sectoral trends at different career stages. Here, predictive margins suggest that early-stage artists typically have lower income levels than the average artist in the total population. But it is again the less advanced, often younger writers that outperform average trends among writers in the sector. Possibly,

[7] Petzold et al. (2020) study the psychological distress, anxiety and depression caused by the Covid-19 pandemic in Germany and find that woman showed overall higher scores than men, in line with other research, e.g. Qiu et al. (2020); Wang et al. (2020).

Covid Economics 82, 23 June 2021: 49-69

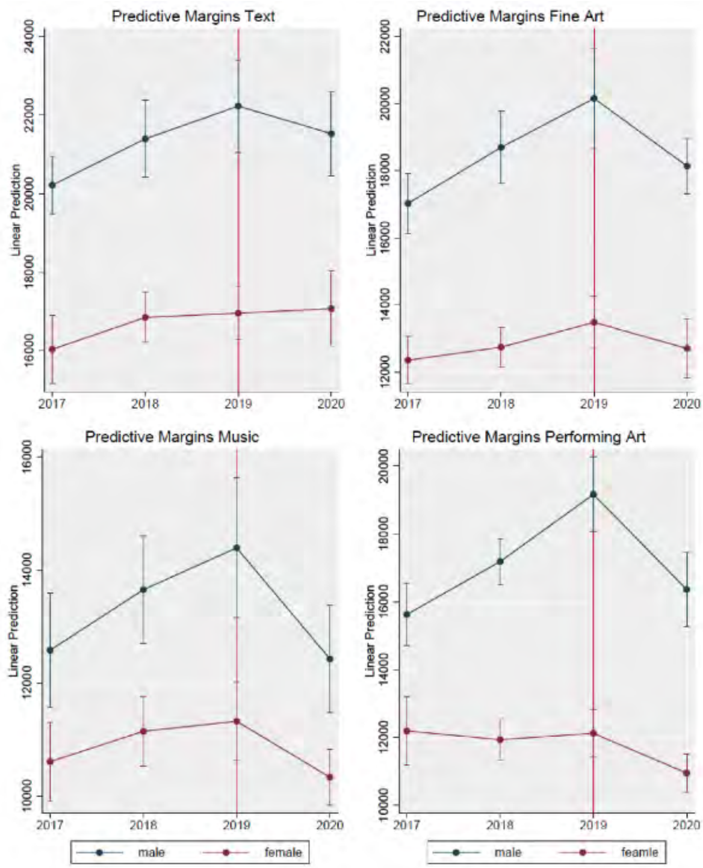


Figure 3: Income by Gender
Note: Panels show predictive margins for each sector and over the observation period. Margins are calculated for women (red) and men (blue).

their work is more reliant on digital sources in the first place and thus may be more resilient to the impact of the crisis. Ultimately, based on the data and the analysis this argument cannot be validated and requires more research. Rather the opposite holds for the Fine Arts. Here, early-stage visual artists are experiencing a more pronounced drop in their incomes than artists that are more advanced in careers.

Covid Economics 82, 23 June 2021: 49-69

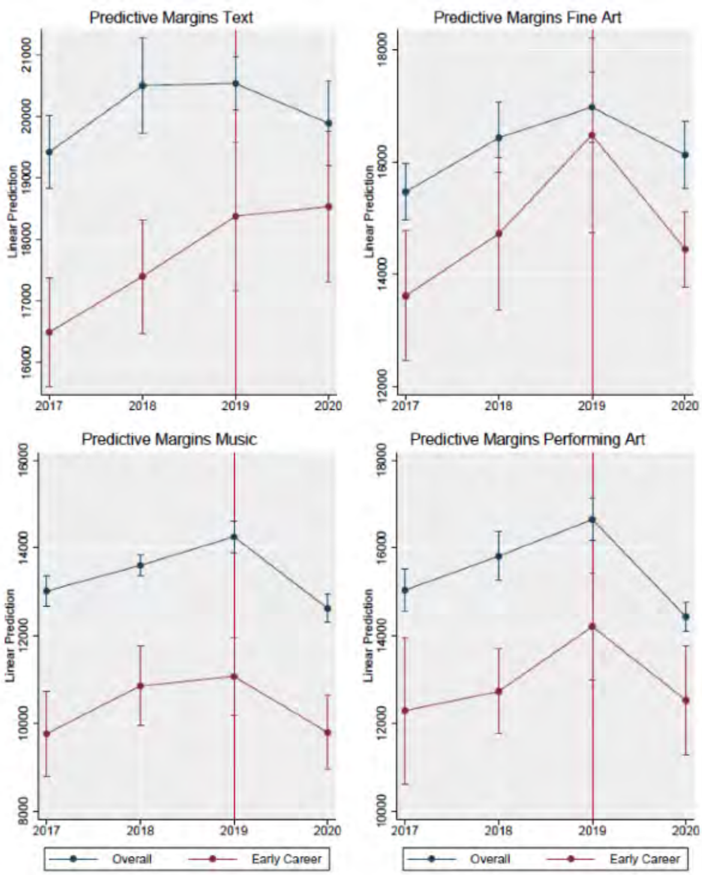


Figure 4: Income by Career Stage

Note: Panels show predictive margins for each sector and over the observation period. Margins are calculated for early stage artists (red) and the total population (blue).

Third, we test if artists located in urban regions see greater income losses due to the pandemic than other regions. This could be due to the fact that some urban agglomerations experienced higher infection rates and saw more restrictive and longer-term NPIs set up. Figure 5 depicts sectoral trends for urban and rural areas in Germany

Covid Economics 82, 23 June 2021: 49-69

depending on population density. Predictive margins indicate that, on average, artists in more populated areas typically earning higher self-employed income. The time trends we can identify, however, do not imply significant variation and differences in the impact of the pandemic.

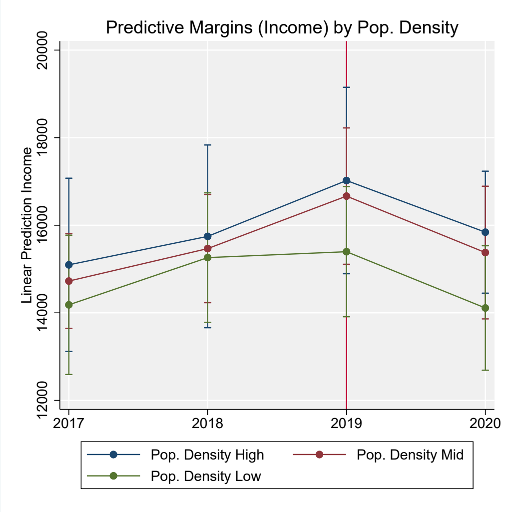


Figure 5: Income by Type of Region

Note: Regions with high/medium/low population density are defined as Pop Density/ km^2 (high) $>$ 1000 $>$ (mid) $>$ 200 $>$ (low). More specifically, the first group includes Berlin, Hamburg, Bremen, the second group includes North Rhine-Westphalia, Saarland, Baden-Wuerttemberg, Hesse, Saxony, Rhineland-Palatinate and the third group the remainder Laender. Source 2017.

4.3 Responsiveness to Specific Public Interventions

Next, we move beyond the binary year treatment used in the baseline model and as outlined in the empirical framework 3. We now allow for a heterogeneous treatment of income groups at the state level based on the estimation model shown in table 1.

CEPR PRESS

However, we introduce an additional measure of "lockdown days" to our models[8], and directly test if income differs for states with fewer lockdown days (i.e. the >25th percentile) and states with longer periods of lockdown (i.e. the <75th percentile). As noted above, this approach does not allow to identify causal estimates, and should be treated and interpreted with great caution as not all sort of state-level variation can be ascribed ex-post to state-level NPI policies.

Coefficient plots in figure 6 help visualize our estimates and serve as a further robustness check, corroborating the baseline results from table 1. In the refined modeling approach, post-2019 income changes continue to show the expected negative sign for all sectors (left-hand side of each panel, coefficient 2020 $Effect$), and the negative gender and early career income effects. On the right-hand side of each panel, we visualize estimates for the interaction terms when accounting for state-level differences in the number of lockdown days. Results in the first panel show that the relative income impact in states with fewer days of lockdown (2020 $\times$ >25th percentile) is slightly higher among writers, fine and performing artists when compared to the same artist groups in other states. Income effects for fine artists are weakly negative under such an intervention which deepens the overall negative impact on their income. For states with longer lockdowns (2020 $\times$ >75th percentile), estimates in the second panel in figure 6 provide evidence on a considerable negative effect, in particular among writers and musicians, whereas in the case of performing artists and fine artists, the overall negative impact on income is reduced with more days of lockdown.

As we cannot detect statistically significant differences on the state-level for the in-

[8]For further details on NPI measures, please see the data section2. We also test income differences resulting from earlier mass-gathering restrictions, pre-school-closing and social distancing rules (not reported here).

Covid Economics 82, 23 June 2021: 49-69

CEPR PRESS

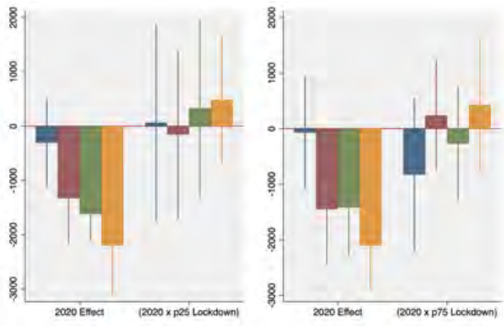


Figure 6: NPI Lockdown Days Coefficients
Note: This figure plots the coefficient size and 90-CI of the interaction term result of interest, based on calculations described in section 3. Writing (blue), fine art (red), music (green) and performing art (yellow), from left to right. Each panel plots the *Post* 2020 Effect on the left-hand-side and the first (second) panel on the right shows the ($2020 \times p25LockdownDays$) interaction term and the ($2020 \times p75LockdownDays$) interaction term, respectively.

teraction coefficients shown in figure 6, this indicate that income losses are not directly correlated and do not systematically vary with specific interventions implemented in each state. Their economic significance is nevertheless important to highlight. Arguably, this sort of heterogeneity in the income effects can help further inform policies and it can be considered an area worth exploring in additional, Covid-19 related economic research.

5 Conclusive Remarks

Admittedly, our findings must be considered preliminary as, based on the data available to us at the end of May 2021, the Covid-19 impact is restricted to 2020 waves and public interventions only, and so we cannot account for 2021 pandemic effects yet. At the same time, we trust that our income estimates do not suffer from funding bias, as much of the financial support granted and distributed from public and private sources

Covid Economics 82, 23 June 2021: 49-69

to artists in Germany will only become visible with the 2021 data.

We provide quantitative evidence that, in the course of this first year, creative and cultural sectors have been impacted differently, with musicians and performing artists experiencing the biggest relative losses in self-employed income. Net income losses range between 2 to 13 percent depending on the sector. Moreover, with writers being the notable exceptions, the crisis does not seem to substantially change the existing gender income gap across sectors nor do income differences at different career stages disappear, all predating the pandemic outbreak.

Finally, tentative analysis that goes beyond mere pandemic-year effects further shows that income levels correlate with the specific way non-pharmaceutical public interventions are being implemented. Here, the quality of interventions affect sectors and groups of artists differently, and so, for example, income from writers seems more sensitive to more days of lockdown than income generated in the performing arts sector, even though these effects are not always statistically significant. This may allow policy makers to more holistically and ex ante assess the choice of public measures in a future pandemic beyond their health goals, and also better target public funding to the most vulnerable groups ex post.

Covid Economics 82, 23 June 2021: 49-69

Covid Economics 82, 23 June 2021: 49-69

References

Aravindakshan, A., Boehnke, J., Gholami, E., and Nayak, A. (2020). Preparing for a future COVID-19 wave: insights and limitations from a data-driven evaluation of non-pharmaceutical interventions in germany. *Scientific Reports*, 10(1):20084.

Baldwin, R. and Di Mauro, B. W. (2020). Economics in the time of covid-19: A new ebook. *VOX CEPR Policy Portal*.

Buchholz, L., Fine, G., and Wohl, H. (2020). Art markets in crisis: how personal bonds and market subcultures mediate the effects of covid-19. *American journal of cultural sociology*, 8:1–15.

Cheng, C., Barceló, J., Hartnett, A. S., Kubinec, R., and Messerschmidt, L. (2020). COVID-19 Government Response Event Dataset (CoronaNet v.1.0). *Nature Human Behaviour*, 4(7):756–768.

Cuntz, A. (2018). *Creators' Income Situation in the Digital Age*, volume 49. WIPO.

Cusmano, L. and Raes, S. (2020). Coronavirus (covid-19): Sme policy responses.

De Vet, J. M., Nigohosyan, D., Ferrer, J., Gross, A.-K., Kuehl, S., and Flickenschild, M. (2021). Impacts of the covid-19 pandemic on eu industries.

Jacobs, R., Finneran, M., and Quintanilla D'Acosta, T. (2020). Dancing toward the light in the dark: Covid-19 changes and reflections on normal from australia, ireland and mexico. *Arts Education Policy Review*, pages 1–10.

Karaca-Mandic, P., Norton, E. C., and Dowd, B. (2012). Interaction terms in nonlinear models. *Health services research*, 47(1pt1):255–274.

Kretschmer, M. (2005). Artists' earnings and copyright: A review of british and german music industry data in the context of digital technologies. *First Monday*, 10(1).

Perraillon, M. C. (2021). Marcelo coca perraillon. url: https://clas.ucdenver.edu/marcelo-perraillon/teaching/health-services-research-methods-i-hsmp-7607. retrieved may 2021.

Petzold, M. B., Bendau, A., Plag, J., Pyrkosch, L., Mascarell Maricic, L., Betzler, F., Rogoll, J., Große, J., and Ströhle, A. (2020). Risk, resilience, psychological distress, and anxiety at the beginning of the covid-19 pandemic in germany. *Brain and behavior*, 10(9):e01745.

Qiu, J., Shen, B., Zhao, M., Wang, Z., Xie, B., and Xu, Y. (2020). A nationwide survey of psychological distress among chinese people in the covid-19 epidemic: implications and policy recommendations. *General psychiatry*, 33(2).

von Bismarck-Osten, C., Borusyak, K., and Schönberg, U. (2020). *The role of schools in transmission of the SARS-CoV-2 virus: Quasi-experimental evidence from Germany.* Number 882. Ruhr Economic Papers.

Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., McIntyre, R. S., Choo, F. N., Tran, B., Ho, R., Sharma, V. K., et al. (2020). A longitudinal study on the mental health of general population during the covid-19 epidemic in china. *Brain, behavior, and immunity*, 87:40–48.

Xue, B. and McMunn, A. (2021). Gender differences in unpaid care work and psychological distress in the uk covid-19 lockdown. *PloS one*, 16(3):e0247959.

Covid Economics 82, 23 June 2021: 49-69

Covid Economics **Issue 82, 23 June 2021**

COVID-19, employment, and gender: Evidence from Nigeria

Marup Hossain[1] and Md Amzad Hossain[2]

Date submitted: 8 June 2021; Date accepted: 17 June 2021

The COVID-19 pandemic-driven economic downturn can have substantial implications for the gender gap in the labor market in developing countries, where women are already worse-off in job participation and earnings than men. Using multiple rounds of individual-level survey data before and after the pandemic and incorporating a difference-in-differences design, we show that overall employment has reduced more for women than men in Nigeria. Women also experienced a larger shift from business employment to farm-based employments. Thus, in addition to causing longer-term unemployment for women, the COVID-19 pandemic may further aggravate women's economic condition to the extent the labor market returns in farming activities are lower than that of business activities.

Covid Economics 82, 23 June 2021: 70-98

1 Economist, Research and Impact Assessment Division, International Fund for Agricultural Development.
2 Assistant professor, Department of Economics, University of Dhaka and Ph.D. candidate Department of Economics, University of Virginia.

1 Introduction

The COVID-19 pandemic has affected the lives and livelihoods of millions around the globe. The developing and least developed countries, in particular, are likely to face stronger and long-lasting negative effects of the COVID-19 pandemic compared to the developed countries because of lack of protection and contingency measures against COVID-19 as well as slow economic recovery rate (ILO, 2021). For instance, studies show that COVID-19 has reduced employment between 5 to 49% and income between 8 to 87% in developing countries (Egger et al., 2021; Khamis et al., 2021). Unlike other recessions (e.g., global financial crisis), COVID-19 has affected women disproportionately in developed countries (Albanesi and Kim, 2021; Alon et al., 2021; Bluedorn et al., 2021). However, little is known about the gender-wise effect of COVID-19 in developing countries, where women are already worse-off in participation, job types, and earning gap compared to men (Jayachandran, 2015).

A few factors might reshape the effects of the COVID-19 pandemic on women in developing and least-developed country context. On the one hand, it might negatively affect women for several reasons. First, women hold temporary jobs more that are more likely to be terminated during a shock (Petrongolo, 2004). Second, pre-existing less access to savings and credit for women accompanied with loss of income (especially as remittances) during the pandemic can widen unequal access to inputs markets that may push more women out of self-employment (World Bank, 2020). Finally, in developing countries where property rights are not well defined, women may lose property rights when their relatives return home from urban areas because of the pandemic and create more competition in acquiring family property (World Bank, 2020). On the other hand, the agriculture sector usually hosts a large fraction of women and, importantly, works as a buffer during shocks –it might favor women in continuing their jobs or entering as an alternative employment option during the pandemic (World Bank, 2021; Christiaensen and Demery, 2017).

Earlier literature investigating the gendered effect of the COVID-19 pandemic fo-

Covid Economics 82, 23 June 2021: 70-98

cuses on developed countries and finds mixed results. For instance, Hupkau and Petrongolo (2020) and Dang and Nguyen (2021) find that the pandemic affected female's job market prospects more than the males in the United States and the United Kingdom, respectively, while Adam-Prassl et al. (2020) document no such difference in Germany. However, to the best of our knowledge, no study systematically explores the pandemic's causal effect in developing countries. We contribute to the literature by providing the first causal estimates of COVID-19 shock on gender inequality in employment outcomes in a developing country context. Since gender inequality and discrimination are more prevalent in developing countries (Jayachandran, 2015), our study fills a critical gap in the literature.

More specifically, this study systematically examines the gender-wise effects COVID-19 pandemic on employment and job composition in a developing country context. We link four waves of individual-level pre-COVID data from the Nigerian Living Standard Survey (LSMS) and ten waves of individual-level post-covid data from COVID-19 National Longitudinal Phone Survey (NLPS).[1] We employ a difference-in-difference approach to identify the causal impact of the COVID-19 pandemic on both overall and gender-wise employment. Our results show that the likelihood of being employed dropped by 13 percent in the post-epidemic period. Women employment reduced about 8 percent more than males in the post-epidemic period. We also find a change in the composition of jobs after the shock – females are more likely to engage in farming activities (52 percent more) and less likely to engage in business activities (44 percent less) than males.

Our findings are consistent over different regression specifications. We also show that results are consistent with the parallel trend assumption, i.e., in the absence of any shock, the difference in outcomes between males and females would be constant over time. Further, applying the event study techniques, we show that the change in job composition remains persistent over time. The duration of unemployment is also longer for females compared to

[1]See Online Appendix Table A.1 for the timing of different rounds of surveys.

Covid Economics 82, 23 June 2021: 70-98

males. All the results indicate that females not only experienced a significant reduction in overall employment compared to males but also went through a larger shift from business to farm employment.

Our findings have serious gender distributional consequences. For instance, We find that women are more likely to be unemployed for longer than men. Earlier literature report that as spells of unemployment prolong, women find it more difficult to reintegrate into the labor market (Alon et al., 2020), and the probability of returning to jobs reduces more for women (New York Times, 2020). Therefore, prolonged unemployment might further dampen the future job prospects of women. Besides, we find that females are more likely to shift from business employment to farm employment. However, agriculture in Africa is characterized by low productivity due to inadequate provision of inputs and extension, and lack of land rights and storage system, which prompted many people to shift from agricultural activity to other sectors (i.e., informal sector) (Barrett et al., 2017; Beegle and Bundervoet, 2019). To make things worse, the agricultural productivity of women-managed land in the African countries are much lower due to the additional constraints that they face compared to men (Mukasa and Salami, 2015; Smith et al., 2015). Therefore, to the extent, labor market returns of farming activities are lower than that of business activities, the COVID-19 pandemic may aggravate the household and individual welfare of females.

The rest of the paper is organized as follows: Section 2 discusses the COVID-19 situation in Nigeria, Section 3 describes the data, and Section 4 discusses the empirical strategy. In Section 5, we present our main results as well as robustness and sensitivity tests. Section 6 concludes our paper.

2 COVID-19 in Nigeria

Nigeria has been affected by the COVID-19 pandemic like many other countries worldwide, with about 164,633 positive cases and 2,061 deaths until April 2021 (NCDC, 2021).[2] Nigeria's

[2]See Online Appendix Figure A.1 to have a visual of COVID-19 spread in Nigeria over time.

CEPR PRESS

federal and state governments have taken various actions such as mandatory use of facemask, physical distancing, avoidance of public gathering, non-essential travel restriction, including a five-week lockdown in Abuja, Lagos, and Ogun States from March 30 to May 03, 2020. However, the recent second wave of COVID-19 cases, which surpassed the first wave in terms of both infection rate and death toll, strongly signals that the economic recovery act will be challenging. Khamis et al. (2021) report that the pandemic-induced work stoppages are about 50% Nigeria—one of the hardest-hit countries in Sub-Saharan Africa. About 80% households also faced some degree of income loss due to the pandemic Khamis et al. (2021). At the national level, the country has just marked its first positive GDP growth (0.1%) in quarter four after two consecutive negative growth rates (3.6% in Q3 and 6.1 % in Q2 of 2020) in the previous two quarters (NBS, 2021).

3 Data

We draw post-COVID data from the COVID-19 National Longitudinal Phone Survey (NLPS) 2020, administered by the National Bureau of Statistics (NBS) of Nigeria and the World Bank. The COVID-19 NLPS collects monthly household-level data on various topics, including the job status of the primary respondents. A total of 10 waves of data are publicly available, starting from April 2020 to February 2021. The NLPS 2020 started with 1,950 samples in the first wave. Some households were dropped in different waves primarily because of refusal from the respondents to be interviewed, resulting in 1,497 households that have been interviewed in all ten waves. The sample of the NLPS is drawn from the fourth round of the General Household Survey—Panel (GHS-Panel) survey (2018-2019), which the NBS and World Bank also administered.

We use the GHS-Panel survey to draw pre-COVID data. We match the GHS-Panel survey data with the NLPS 2020 in two ways. First, we match respondents of the NLPS 2020 with the fourth wave of the GHS-Panel survey and get 1,200 matched samples. This approach provides us 12 waves of data, ten waves of the NLPS 2020, and two waves of

Covid Economics 82, 23 June 2021: 70-98

the GHS-Panel fourth round. Second, although the first approach generates 1,200 unique samples, it generates only two waves of pre-COVID data, limiting the scope of examining the pre-COVID parallel trend between groups. We could not go beyond the fourth round because the individual level matching yields a small sample (about 300).[3] However, about 80% of the respondents in the NLPS 2020 survey are household heads, allowing us to draw samples at the household head level. Therefore, we match the sample at the household head level for both NLPS and GHS-Panel surveys. Since the second approach generates a much bigger sample and allows us to draw a sample from the previous rounds of the GHS-Panel survey, we use this sample as our primary data. However, we also report results using the individual-level matched data to show whether our results are sensitive to different data sets.

3.1 Outcome variables

Our main outcome variable is whether a person was involved in any income-generating activities in the preceding week of the survey interview. The NLPS 2020 survey asked the respondent whether they participated in an activity in the past seven days with a follow-up question on types of activity. Based on the recorded types of activity, we measure three additional dummy outcome variables: farming (takes a value of one if the respondent worked in a family farm growing crops, raising livestock, or fishing), business (takes a value of one if the respondent worked in own business or in a business operated by a household or family member), and wage/service (takes a value of one if the respondent worked for someone who is not a household member). In the GHS-Panel survey, the question structures were a little different: it collected information about the three job categories (farming, business, and wage/service) separately for the same recall period. From these three individual indicators, we measure the aggregate indicator variable (involvement in any income-generating activities) that takes a value of one if the person worked in at least one activity.

[3]In the fourth round of the GHS-Panel survey, the LSMS–ISA team added 3,600 new households in the latest round of the survey by dropping an equal number of existing households in round one to three, leaving only about 1,500-panel households

3.2 Summary Statistics

About 80% of our sample were employed either in farming, business, or wage/service activity before the COVID-19 pandemic.[4] During the early period of the pandemic (March to May 2020), overall employment plummeted to 43%, so did sector-wise employments. The overall employment rate swung back to pre-COVID rates after that period, although the farming and business activities never recovered fully. As shown in Figure 1, overall employment rates are lower for females than males in both pre-and post-COVID periods. One striking result stands out from the farming employment distribution over time: the female employment rate surpassed the male employment rate during the post-COVID periods. In contrast, the business employment rate declined for females compared to males during the post-COVID periods.

4 Empirical Approach

We employ a difference-in-difference approach to identify the causal impact of COVID-19 pandemic on overall and gender-wise employment. First, to estimate the effect on overall employment, we compare regions with higher COVID intensity to regions with lower COVID intensity before and after the COVID-19 breakout. More precisely, we run the regression of the following form:

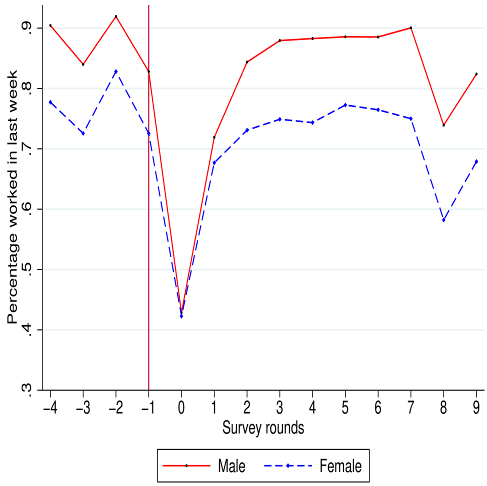
$$\begin{aligned} y_{ihst} = \alpha_{hs} + \beta_t + \gamma_1 HighIntensity_{st} + \gamma_2 Post_t + \gamma_3 HighIntensity_{st} \times Post_t \\ + \delta X_{ihst} + u_{ihst} \end{aligned} \tag{1}$$

where y_{ihst} is the outcome of interest for individual i from household h in state s in survey round t, $HighIntensity_{st}$ is a dummy variable taking a value of 1 if the household is from a state s with more than the median infection rates during the early phase of COVID-19 pandemic, and 0 otherwise, $Post_t$ is an indicator variable taking a value of 1 for the survey rounds after COVID-19 breakout, and 0 otherwise. α_{hs} are the household fixed effects

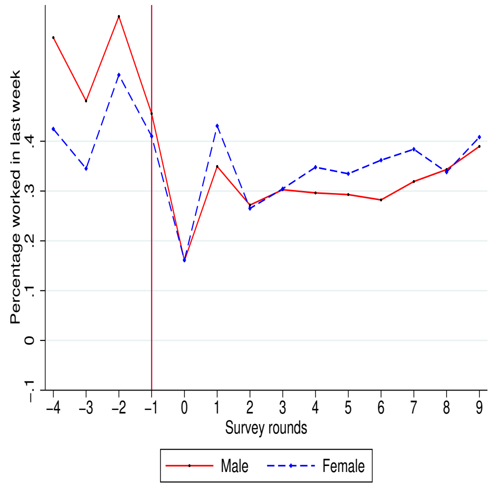
[4]See online Appendix Table A.2 to see the job compositions by gender over different rounds.

Covid Economics 82, 23 June 2021: 70-98

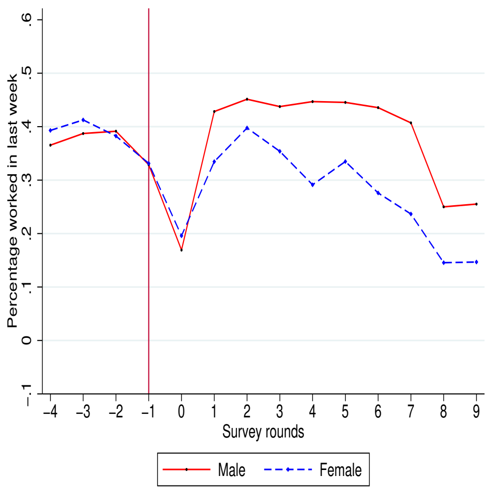
Covid Economics 82, 23 June 2021: 70-98



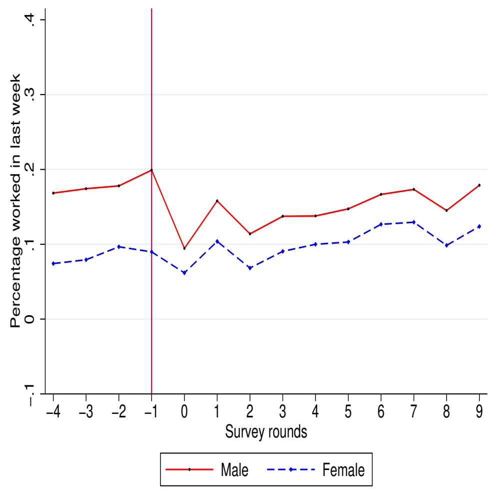
(a) Any work



(b) Farming



(c) Business



(d) Wage/Service

Figure 1: Job Participation Rates Across Rounds by Gender

Notes: Graph shows gender wise work status of the household heads in the preceding week of the different survey rounds. Farming includes self-employment in a household based crops cultivation, other farming tasks, or livestock activity. Business includes household based enterprise, for example, as a trader, shopkeeper, barber, dressmaker, carpenter or taxi driver. Wage/service includes work for a non-household member, fr example, an enterprise, company, the government or any other individual. Finally, any work is a combination of farming, business, and wage/service.

Covid Economics 82, 23 June 2021: 70-98

which would control for all time-invariant household characteristics that affect the outcome variables, and β_t are the survey round fixed effects which would control for survey round specific shocks which are common across all households. X_{ihst} is a vector for other time-varying individual characteristics such as age, marital status, and education level. We are mainly interested in the parameter γ_3, which will show the causal estimates conditional on some assumptions as explained below.

Next, we examine the effects of the pandemic on employment by gender, again through a difference-in-differences framework. To this end, we run a regression of the following form:

$$\begin{aligned} y_{ihst} = \alpha_{hs} + \beta_t + \gamma_1 Female_{ihst} + \gamma_2 Post_t + \gamma_3 Female_{ihst} \times Post_t \\ + \delta X_{ihst} + u_{ihst} \end{aligned} \tag{2}$$

Where $Female_{ihst}$ is a dummy variable taking a value of 1 if the individual is female and 0 otherwise. All other variables are similar to those used in Equation 1. The parameter γ_3 will capture the gender-differential effect of the pandemic on employment: a positive coefficient will imply that females are performing better than the male with respect to employment after the shock, and vice versa.

One important assumption for our identification is the *parallel trend* assumption, i.e., high-intensity and low-intensity regions or gender-wise employment would follow a similar trend in the absence of the COVID-19 shock. While we cannot test the parallel trend assumption directly, we can test one implication of the Parallel trend assumption – in the pre-shock period, time trends in the outcome are the same in treated and control units (i.e., parallel pre-trend). To test for parallel pre-trend, we run the regression of the following form:

$$\begin{aligned} y_{ihst} = \alpha_{hs} + \beta_t + \gamma_{-k} D_{ihst} + \gamma_{-k-1} D_{ihst} + \gamma_{-k-2} D_{ihst} + ... + \gamma_{-1} D_{ihst} + \gamma_1 D_{ihst} \\ + \gamma_2 D_{ihst} + ... + \gamma_l D_{ihst} + \delta X_{ihst} + u_{ihst} \end{aligned} \tag{3}$$

Where D_{ihst} is a dummy variable indicating gender or COVID intensity. We include the interactions of the time dummies and the treatment indicator for all pre- and post-periods and leave out the one interaction for the last pre-treatment period, which serves as the baseline. If the difference between treatment and control group remains the same over time, we would expect γ_{-k}, γ_{-k-1}, ... , γ_{-1} to be insignificant. We test the validity of our identifying assumption, as well as offer several falsification tests.

5 Results

We first look at the effect of the COVID-19 pandemic on the probability of different types of employment. Column (1) of Table 1 shows that the likelihood of being employed in the preceding week of the interview fell by 10 percentage points after the pandemic, a reduction of about 13 percent in the likelihood of employment. The decline was much higher (an additional 5 percentage points) for households from states with higher COVID intensity than those from states with lower COVID intensity.

Table 1: Effect of COVID-19 on the Probability of Employment

	(1) Any work	(2) Farming	(3) Business	(4) Wage/services
COVID intensity	0.0443	-0.0315	-0.198	0.180*
	(0.0373)	(0.155)	(0.175)	(0.105)
Post	-0.101***	-0.213***	-0.0587	-0.0574**
	(0.0225)	(0.0626)	(0.0525)	(0.0235)
COVID intensity X Post	-0.0542*	0.192*	-0.263***	-0.00916
	(0.0287)	(0.0985)	(0.0734)	(0.0231)
Constant	1.076***	0.623***	0.732***	0.0519
	(0.0703)	(0.117)	(0.143)	(0.0702)
Observations	33,034	33,034	33,034	33,034
R-squared	0.502	0.572	0.510	0.659
Endline control mean	0.784	0.269	0.416	0.117

Notes: Data are drawn from the Nigerian Living Standard Survey and are at the household level. Post is a dummy variable taking a value of 1 for survey rounds taking place after the COVID-19 pandemic took place, and 0 otherwise. COVID intensity is a dummy variable taking a value of 1 if the household is from a State where COVID-19 infection rates were above the median value of the State-wise infection rate distribution during March 1st to April 30, 2020. Each regression controls for household and survey round fixed effects. Each regression control for socio-economic controls which includes age, marital status, and education level of household head. Standard errors are clustered at the state level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

We then look at the probability of employment by type of employment – Farming, business, or wage employment. Column (2) of Table 1 suggests that the probability of employment in farming activities and business activities was 19 percentage points higher and 26 percentage points lower, respectively, for high COVID intensity states compared to the lower-intensity states. We do not observe any significant difference in wage or service-based employment between the high and low-intensity states. The results suggest that the pandemic resulted in a shift from business employment to farm employment with a reduction in overall employment.

We also look at how the effect of the pandemic varied over time. To this end, we run regressions using specification 3 and plot the coefficients of the interaction of the COVID intensity dummy with the year dummies. The event study in Figure A.2 shows that the probability of employment fell significantly immediately after the COVID-19 pandemic and the resulting lockdown (Panel A). However, the treatment effect faded over time, and the labor market participation rate returned to the pre-COVID situation within a few months. While the overall employment was almost unchanged in the long run, there was a significant change in the composition of jobs. The probability of employment in business activities significantly fell immediately after the shock and never recovered (Panel C). Households tried to cope by engaging more in farm activities, and, therefore, the probability of employment in farming activities increased following the pandemic and remained the same over time (Panel B).

We then explore the gender-distributional effect of the COVID-19 pandemic on employment (Table 2). It is evident that females were affected more by the pandemic than males – the probability of any employment was 6 percentage points lower for females than males in the post-epidemic period. When we explore the effect of the pandemic on employment by type, we find that females were more likely to engage in farming activities (16 percentage points) – about 52 percent higher than the endline control mean. However, the females were 16 percentage points less likely to engage in business activities than males. Wage and service

Covid Economics 82, 23 June 2021: 70-98

jobs also increased for females by 4 percentage points during the pandemic. Thus we see a higher transition from business activities to agricultural activities for women than men. This should not come as a surprise since the agriculture sector usually hosts a large fraction of women and, importantly, works as a buffer during shocks, and, therefore, it might favor women in continuing their jobs or entering as an alternative employment option during the pandemic (World Bank, 2021; Christiaensen and Demery, 2017).

Table 2: Effect of COVID-19 on the Probability of Employment by Gender

	(1) Any work	(2) Farming	(3) Business	(4) Wage/services
Female	-0.0400	-0.0147	0.0176	-0.0920***
	(0.0533)	(0.0696)	(0.0610)	(0.0301)
Post	-0.115***	-0.144**	-0.164***	-0.0658***
	(0.0221)	(0.0584)	(0.0427)	(0.0210)
Female X Post	-0.0606**	0.155*	-0.164*	0.0437*
	(0.0285)	(0.0796)	(0.0813)	(0.0216)
Constant	1.135***	0.595***	0.645***	0.196***
	(0.0809)	(0.106)	(0.119)	(0.0565)
Observations	33,034	33,034	33,034	33,034
R-squared	0.502	0.570	0.504	0.659
Endline control mean	0.793	0.298	0.372	0.144

Notes: Data are drawn from the Nigerian Living Standard Survey and are at the household level. Post is a dummy variable taking a value of 1 for survey rounds taking place after the COVID-19 pandemic took place, and 0 otherwise. Female is a dummy variable taking a value of 1 if household head is female, and 0 otherwise. Each regression controls for household and survey round fixed effects. Each regression control for socio-economic controls which includes age, marital status, and education level of household head. Standard errors are clustered at the state level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Our results are somewhat similar to those of Canjer et al. (2020), Hupkau and Petrongolo (2020), and Dang and Nguyen (2021) who find that the pandemic affected female's job market prospect more than the males in different developing countries. For instance, Dang and Nguyen (2021) study the effect of the COVID-19 pandemic on gender inequality in the labor market using a representative sample from six countries –China, South Korea, Japan, Italy, the United Kingdom, and the United States– and find that women are 24 percent more likely to lose jobs permanently than men due to the pandemic. While we find a much smaller effect size compared to the study above, we document a note-able change

Covid Economics 82, 23 June 2021: 70-98

in the composition of jobs –a large shift from agricultural activities to business activities for women. Our results, however, are different from Adams-Prassl et al. (2020), which show that women are not more likely to lose jobs than men in Germany.

Next, we examine the gender-distributional effect of COVID-19 shock on employment by the severity of the pandemic. We report the results in Table 3, which reinforces our finding from Table 1 that the high-intensity states experienced a shift from business employment to farm employment after the shock. Table 3 further shows that the shift was much larger for the females. The probability of employment in farming activities for females increased by 38 percentage points in the high-intensity regions after the pandemic. On the contrary, the probability of employment in business activities for females fell by 33 percentage points in the high-intensity states. Overall, female from the high-intensity regions were 4.7 percentage points less likely to be employed compared to the male. More importantly, there has been a change in the composition of jobs for the female – females switched more from business activities to farming activities after the pandemic. As seen in the event study Figure 2, this larger shift from business employment to farm employment for women was not a one-shot event. Rather, the shift persisted over time.

A larger shift from business activities to farm activities for females might have important gender-distributional consequences. Agriculture in Africa is characterized by low productivity due to inadequate provision of inputs and extension, and lack of land rights and storage system, which induced many farm households to shift from agricultural sector to other sectors (i.e., informal sector) (Barrett et al., 2017; Beegle and Bundervoet, 2019). Besides, the agricultural productivity of women-managed land in the African countries are much lower, not necessarily due to the fact that females are less productive, but because of other constraints that they face compared to men (Mukasa and Salami, 2015; Smith et al., 2015). For instance, Jayachandran (2015) cited factors like physical strength or brawn and social norms to explain the gender gap in agriculture. All these evidences suggest that fe-

Table 3: COVID-19 Intensity and the Probability of Employment by Gender

	(1) Any work	(2) Farming	(3) Business	(4) Wage/services
Female	0.00213	0.0623	0.00902	-0.0989***
	(0.0685)	(0.0914)	(0.0807)	(0.0360)
Post	-0.0896***	-0.219***	-0.0476	-0.0604**
	(0.0223)	(0.0630)	(0.0542)	(0.0249)
Female X Post	-0.0421	0.00833	-0.0392	0.0298
	(0.0335)	(0.0819)	(0.0864)	(0.0336)
COVID intensity	0.0566	0.0284	-0.224	0.170
	(0.0340)	(0.159)	(0.185)	(0.102)
Female X COVID intensity	-0.0774	-0.236**	0.0883	0.0146
	(0.0896)	(0.107)	(0.0923)	(0.0514)
Post X COVID intensity	-0.0471	0.138	-0.215***	-0.0119
	(0.0331)	(0.0967)	(0.0735)	(0.0243)
Female X Post X COVID intensity	-0.0477	0.375***	-0.332***	0.0287
	(0.0535)	(0.104)	(0.104)	(0.0419)
Constant	1.105***	0.590***	0.747***	0.113
	(0.0878)	(0.132)	(0.152)	(0.0697)
Observations	33,034	33,034	33,034	33,034
R-squared	0.503	0.576	0.513	0.659
Endline control mean	0.793	0.298	0.372	0.144

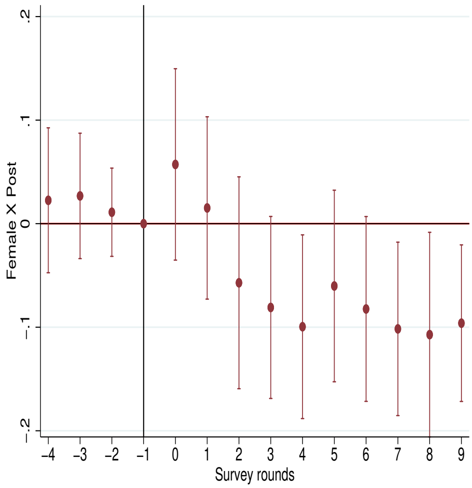
Notes: Data are drawn from the Nigerian Living Standard Survey and are at the household level. Post is a dummy variable taking a value of 1 for survey rounds taking place after the COVID-19 pandemic took place, and 0 otherwise. Female is a dummy variable taking a value of 1 if household head is female, and 0 otherwise. Each regression controls for household and survey round fixed effects. Each regression control for socio-economic controls which includes age, marital status, and education level of household head. Standard errors are clustered at the state level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

males are shifting towards a sector with lower return more than the males in the long-run. This may further aggravate the already existing gender gap in labor market.

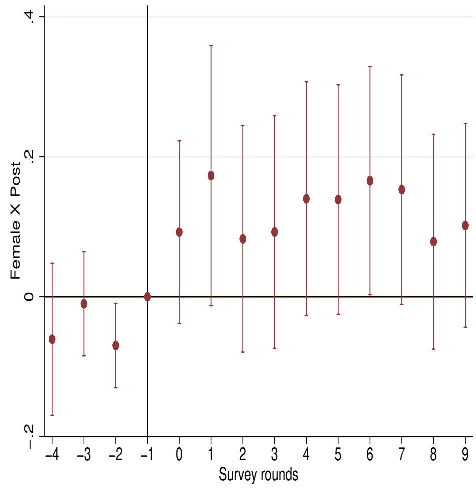
We examine the gender-wise employment effects of COVID-19 further. We find that females remained unemployed for a more extended period compared to males after the initial COVID-19 exposure (Online Appendix Figure A.5). Conditional on being unemployed at least once, females remain unemployed for 3.67 months, on an average, whereas the corresponding number for males is 2.61. This finding has several implications on the future job prospects of women. Earlier literature documents that with longer spells of unemployment, women find it more difficult to reintegrate into the labor market (Alon et al., 2020), and the probability of returning to jobs reduces more for women (New York Times, 2020). Therefore, prolonged unemployment might further dampen the future job prospects of women and increase the gender gap in labor market.

Covid Economics 82, 23 June 2021: 70-98

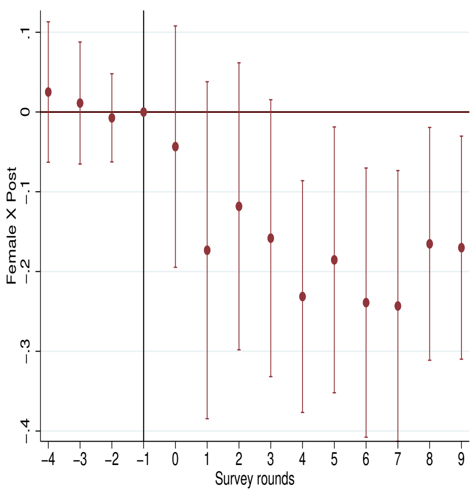
Covid Economics 82, 23 June 2021: 70-98



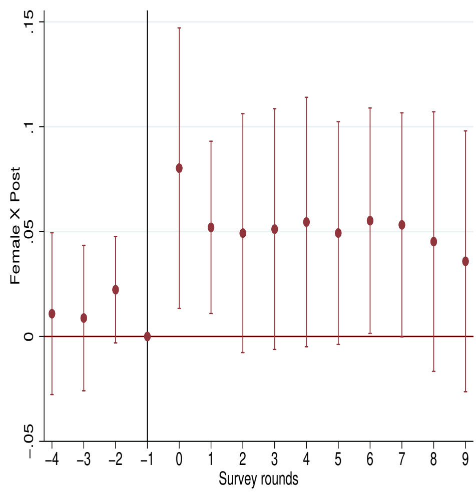
(a) Any work



(b) Farming



(c) Business



(d) Wage/Service

Figure 2: Event study: Effect of COVID-19 on the Probability of Employment by Gender

Notes: Graphs plot the coefficient of the interaction variable $Female \times post$ from the regression specified in equation 2. Solid circles show coefficients and maroon lines show corresponding 95% confidence intervals. Regression controls for household and survey round fixed effects, as well as, socio-economic controls such as age, marital status, and education level of household head. Standard errors are clustered at the state level.

CEPR PRESS

5.1 Sensitivity and Robustness

We first test for parallel trend assumption, i.e., in the absence of any shock, the difference in outcomes between males and females would be constant over time. Two pieces of evidence suggest that parallel trend assumption hold in our settings. First, visual inspection of Figure 1 suggests that both males and females exhibited a similar kind of labor market participation trend before the shock. Second, the event study plots suggest that the difference in trend between high-intensity regions and low-intensity regions (Online Appendix Figure A.2) or male and female (Figure 2) before the start of the pandemic is not statistically significant. Thus we cannot reject the null hypothesis of parallel trend.

In addition, we check our results using a second sample where we match individuals instead of household heads over different survey waves and controlling for individual fixed effects.[5] We find that the results are very similar to the ones we got using our primary specifications (See Online Appendix Table A.3 for the summary statistics of this sample and Online Appendix Tables A.4 - A.6 for the results)

6 Conclusion

In this paper, we made the first attempt to explore the gendered effect of the COVID-19 pandemic on employment in a developing country context. We use a difference-in-differences technique to show that regions with higher COVID intensity experienced a shift from business activities to farm employment. Women not only experienced a significant reduction in overall employment compared to males but also went through a larger shift from business to farm employment. We also find that the gendered effect of the pandemic persists over time – women who lost jobs are less likely to rejoin the labor market when the labor market conditions improve.

[5]Note that this allows us to use only two pre-period survey rounds, as the earlier rounds had much less sample size and, therefore, matching individuals across the surveys results in significant reduction in the sample size.

Covid Economics 82, 23 June 2021: 70-98

CEPR PRESS

Understanding how the pandemic affected labor markets in the developing world is crucial as governments and other actors continue to develop responses. Therefore, our findings can help inform the policy makers to formulate appropriate short-term and medium-term policy responses aiming at ameliorating the impacts of COVID-19. Our results suggest that the pandemic has widened gender inequality in labor market. To address this growing inequality in the coming months, we need to see strong policies to support women, in the absence of which female unemployment might mount, worsening gender inequality.

Covid Economics 82, 23 June 2021: 70-98

References

Adams-Prassl, A., Boneva, T., Golin, M., and Rauh, C. (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics*, 189:104245.

Albanesi, S. and Kim, J. (2021). The Gendered Impact of the COVID-19 Recession on the US Labor Market. Technical report, National Bureau of Economic Research.

Alon, T., Coskun, S., Doepke, M., Koll, D., and Tertilt, M. (2021). From Mancession to Shecession: Women's Employment in Regular and Pandemic Recessions. Technical report, National Bureau of Economic Research.

Alon, T. M., Doepke, M., Olmstead-Rumsey, J., and Tertilt, M. (2020). The impact of COVID-19 on gender equality. Technical report, National Bureau of economic research.

Barrett, C. B., Christiaensen, L., Sheahan, M., and Shimeles, A. (2017). *On the structural transformation of rural Africa*. The World Bank.

Beegle, K. and Bundervoet, T. (2019). Moving to jobs off the farm.

Bluedorn, J., Caselli, F., Hansen, N.-J., Shibata, I., and Tavares, M. M. (2021). Gender and Employment in the COVID-19 Recession: Evidence on "She-cessions".

Christiaensen, L. and Demery, L. (2017). *Agriculture in Africa: Telling myths from facts*. The World Bank.

Dang, H.-A. H. and Nguyen, C. V. (2021). Gender inequality during the COVID-19 pandemic: Income, expenditure, savings, and job loss. *World Development*, 140:105296.

Egger, D., Miguel, E., Warren, S. S., Shenoy, A., Collins, E., Karlan, D., Parkerson, D., Mobarak, A. M., Fink, G., Udry, C., et al. (2021). Falling living standards during the

COVID-19 crisis: Quantitative evidence from nine developing countries. *Science advances*, 7(6):eabe0997.

Hupkau, C. and Petrongolo, B. (2020). Work, care and gender during the Covid-19 crisis. *Fiscal studies*, 41(3):623–651.

ILO (2021). COVID-19 - Tackling the jobs crisis in the Least Developed Countries. Technical report.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *economics*, 7(1):63–88.

Khamis, M., Prinz, D., Newhouse, D., Palacios-Lopez, A., Pape, U., and Weber, M. (2021). The Early Labor Market Impacts of COVID-19 in Developing Countries.

Mukasa, A. N. and Salami, A. O. (2015). *Gender productivity differentials among smallholder farmers in Africa: A cross-country comparison.* African Development Bank Abidjan.

NBS (2021). Nigerian Gross Domestic Product Report. Technical report.

NCDC (2021). COVID-19 NIGERIA. Accessed: 2021-04-24.

New York Times (2020). Why Did Hundreds of Thousands of Women Drop Out of the Work Force? Accessed: 2021-04-28.

Petrongolo, B. (2004). Gender segregation in employment contracts. *Journal of the European Economic Association*, 2(2-3):331–345.

Smith, D., Torkelsson, A., and Westman, M. (2015). The cost of the gender gap in agricultural productivity in malawi, tanzania, and uganda. *New York: UN Women.*

World Bank (2020). COVID-19: A pivotal moment to support women farmers. Accessed: 2021-04-28.

Covid Economics 82, 23 June 2021: 70-98

World Bank (2021). Agriculture as a buffer in COVID-19 crisis: Evidence from five Sub-Saharan African countries. Accessed: 2021-05-16.

Covid Economics 82, 23 June 2021: 70-98

CEPR PRESS

Online Appendix

Covid Economics 82, 23 June 2021: 70-98

Covid Economics 82, 23 June 2021: 70-98

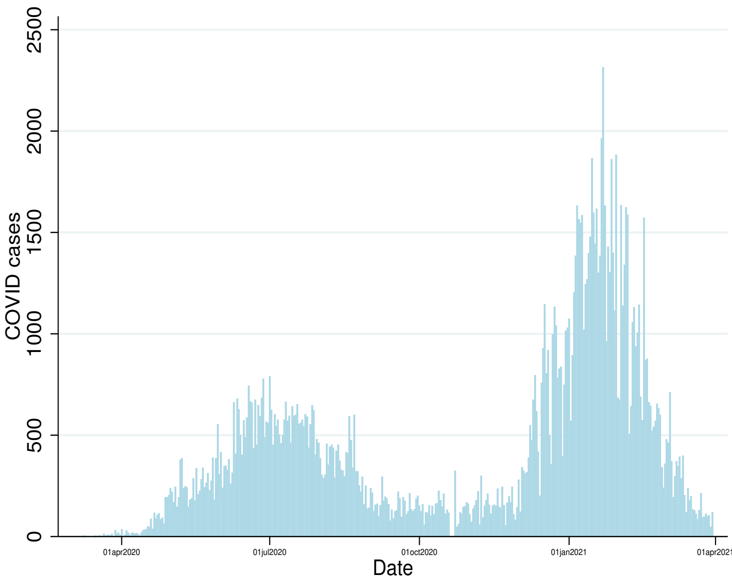


Figure A.1: COVID Contamination over Time in Nigeria

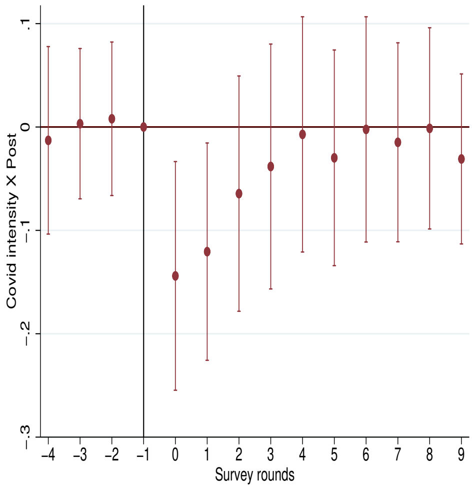
Notes: Graph shows daily infection rate of COVID-19 cases in Nigeria. Data are drawn from the Nigeria Centre for Disease Control through the Humanitarian Data Exchange Portal.

Table A.1: Survey timing

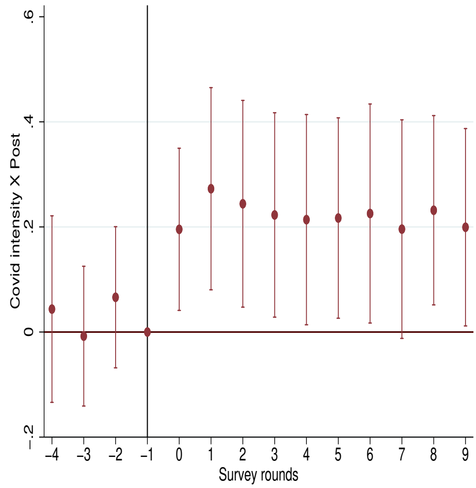
Round	(1) Survey Start	(2) Survey End
Round -4 (Post-planting)	08/2015	10/2015
Round -3 (Post-harvest)	02/2016	04/2016
Round -2 (Post-planting)	07/2018	09/2018
Round -1 (Post-harvest)	01/2019	02/2019
Round 1	4/20/2020	5/11/2020
Round 2	6/2/2020	6/16/2020
Round 3	7/6/2020	7/20/2020
Round 4	8/9/2020	8/24/2020
Round 5	9/7/2020	9/21/2020
Round 6	10/9/2020	10/24/2020
Round 7	11/7/2020	11/23/2020
Round 8	12/5/2020	12/21/2020
Round 9	1/9/2021	1/25/2021
Round 10	2/6/2021	2/22/2021

Notes: Information are dwarn from the World Bank website. Available here.

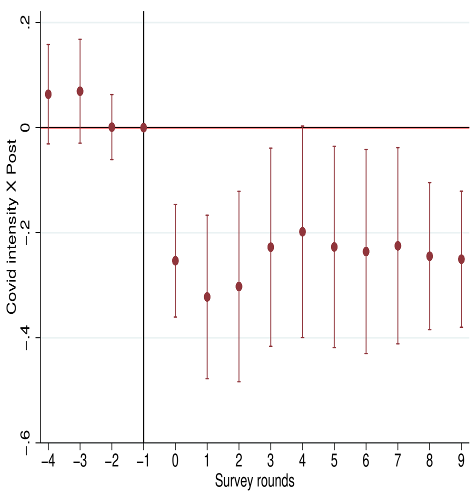
Covid Economics 82, 23 June 2021: 70-98



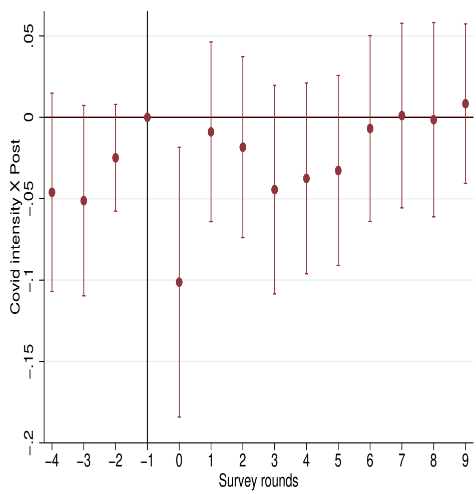
(a) Any work



(b) Farming



(c) Business

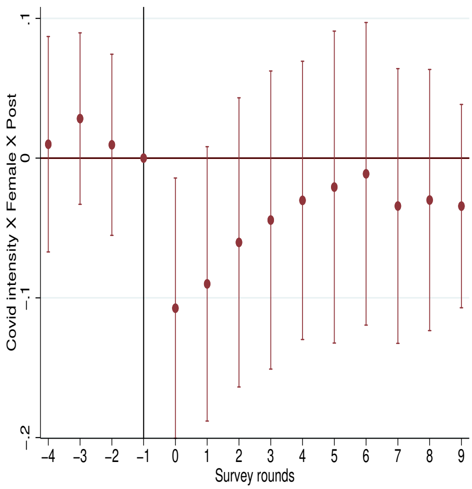


(d) Wage/Service

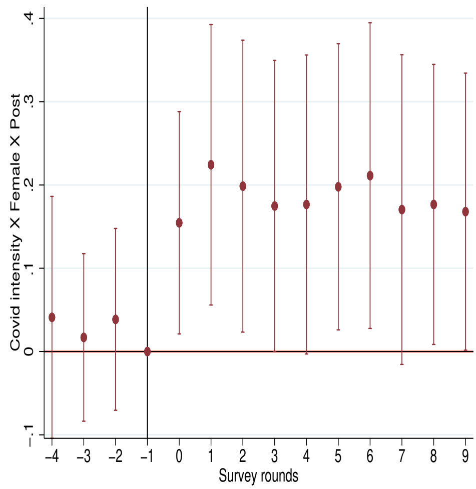
Figure A.2: Event Study: Effect of COVID-19 on Probability of Employment Over Time

Notes: Graphs plot the coefficient of the interaction variable $COVIDIntensity \times post$ from the regression specified in equation 1. Solid circles show coefficients and maroon lines show corresponding 95% confidence intervals. Regression controls for household and survey round fixed effects, as well as, socio-economic controls such as age, marital status, and education level of household head. Standard errors are clustered at the state level.

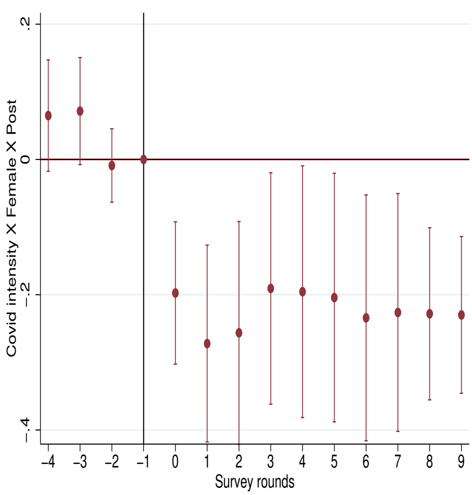
Covid Economics 82, 23 June 2021: 70-98



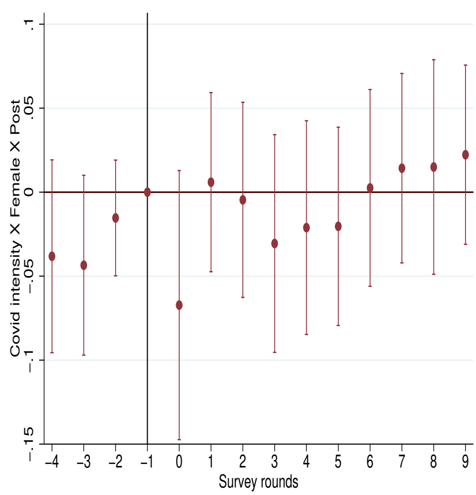
(a) Any work



(b) Farming



(c) Business

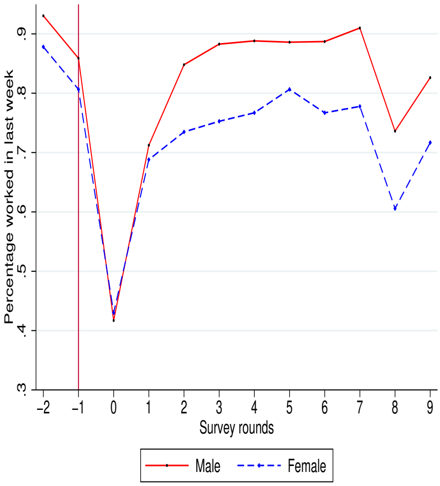


(d) Wage/Service

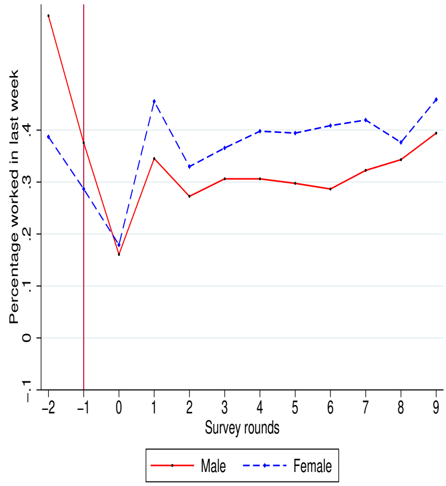
Figure A.3: Event study: Effect of COVID-19 on Gender Gap in Employment by the Severity of the Pandemic

Notes: Graphs plot the coefficient of the interaction variable $COVIDIntensity \times post \times Female$ from the regression of probability of employment on COVID intensity dummy, Post dummy, Female dummy, and their interactions. Solid circles show coefficients and maroon lines show corresponding 95% confidence intervals. Regression controls for household and survey round fixed effects, as well as, socio-economic controls such as age, marital status, and education level of household head. Standard errors are clustered at the state level.

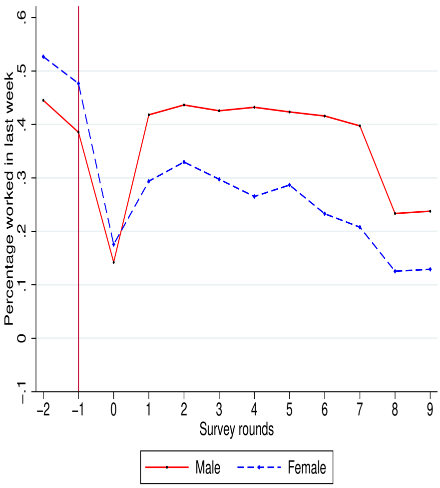
Covid Economics 82, 23 June 2021: 70-98



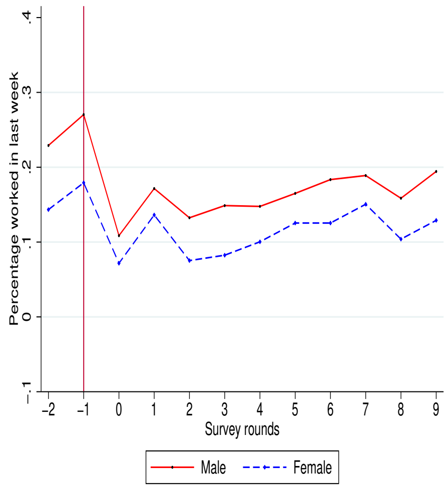
(a) Any work



(b) Farming



(c) Business



(d) Wage/Service

Figure A.4: Job Participation Rates Across Rounds by Gender

Notes: Graph shows gender wise work status of the survey respondent in the preceding week of the different survey rounds. Data are drawn at the survey respondent level; only matched respondents are used in the analyses. Number of of observation is 1,200 in each rounds. Farming includes self-employment in a household based crops cultivation, other farming tasks, or livestock activity. Business includes household based enterprise, for example, as a trader, shopkeeper, barber, dressmaker, carpenter or taxi driver. Wage/service includes work for a non-household member, fr example, an enterprise, company, the government or any other individual. Finally, any work is a combination of farming, business, and wage/service.

Covid Economics 82, 23 June 2021: 70-98

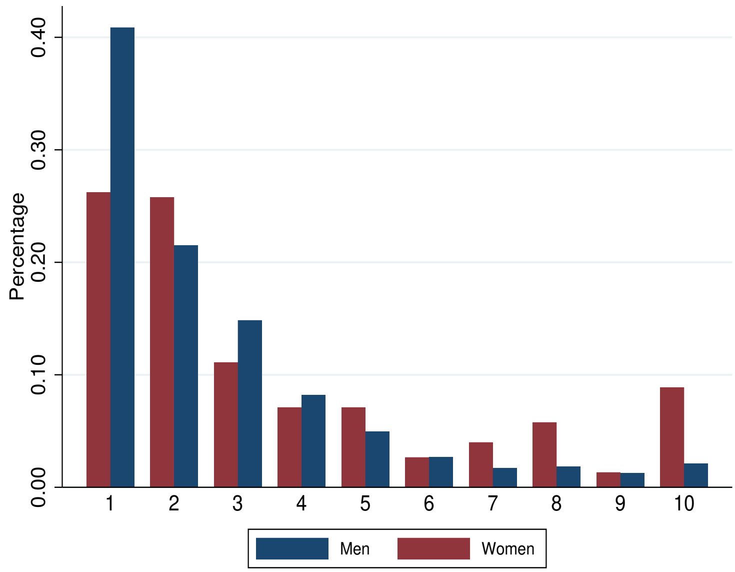


Figure A.5: Unemployment periods by Gender

Notes: Graph shows number of periods individuals were not working by Gender. Only Post-COVID rounds data are used. Maximum number of periods can be 10.

CEPR PRESS

Table A.2: Summary Statistics over Survey Rounds and Gender

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Round	Any work		Farming		Business		Wage/services		Observation	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
-4	0.78	0.90	0.42	0.61	0.39	0.37	0.07	0.17	916	3,675
-3	0.73	0.84	0.35	0.48	0.41	0.39	0.08	0.17	933	3,649
-2	0.83	0.92	0.53	0.65	0.38	0.39	0.10	0.18	983	4,067
-1	0.73	0.83	0.41	0.46	0.33	0.33	0.09	0.20	1,002	3,978
1	0.42	0.43	0.16	0.16	0.20	0.17	0.06	0.09	291	1,312
2	0.68	0.72	0.43	0.35	0.33	0.43	0.10	0.16	260	1,228
3	0.73	0.84	0.27	0.27	0.40	0.45	0.07	0.11	249	1,203
4	0.75	0.88	0.30	0.30	0.35	0.44	0.09	0.14	243	1,186
5	0.74	0.88	0.35	0.30	0.29	0.45	0.10	0.14	230	1,168
6	0.77	0.89	0.33	0.29	0.33	0.45	0.10	0.15	233	1,154
7	0.76	0.89	0.36	0.28	0.28	0.44	0.13	0.17	221	1,134
8	0.75	0.90	0.38	0.32	0.24	0.41	0.13	0.17	224	1,125
9	0.58	0.74	0.34	0.34	0.15	0.25	0.10	0.15	213	1,116
10	0.68	0.82	0.41	0.39	0.15	0.26	0.12	0.18	218	1,101

Notes: Graph shows gender wise work status of the survey respondent in the preceding week of the different survey rounds. Farming includes self-employment in a household based crops cultivation, other farming tasks, or livestock activity. Business includes household based enterprise, for example, as a trader, shopkeeper, barber, dressmaker, carpenter or taxi driver. Wage/service includes work for a non-household member, fr example, an enterprise, company, the government or any other individual. Finally, any work is a combination of farming, business, and wage/service.

Covid Economics 82, 23 June 2021: 70-98

Table A.3: Summary Statistics over Survey Rounds and Gender

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Round	Any work		Farming		Business		Wage/services		Observation	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
-2	0.88	0.93	0.39	0.62	0.53	0.45	0.14	0.23	279	921
-1	0.81	0.86	0.29	0.38	0.48	0.39	0.18	0.27	279	921
1	0.43	0.42	0.18	0.16	0.18	0.14	0.07	0.11	279	921
2	0.69	0.71	0.46	0.35	0.29	0.42	0.14	0.17	279	921
3	0.73	0.85	0.33	0.27	0.33	0.44	0.08	0.13	279	921
4	0.75	0.88	0.37	0.31	0.30	0.43	0.08	0.15	279	921
5	0.77	0.89	0.40	0.31	0.27	0.43	0.10	0.15	279	921
6	0.81	0.89	0.39	0.30	0.29	0.42	0.13	0.17	279	921
7	0.77	0.89	0.41	0.29	0.23	0.42	0.13	0.18	279	921
8	0.78	0.91	0.42	0.32	0.21	0.40	0.15	0.19	279	921
9	0.61	0.74	0.38	0.34	0.13	0.23	0.10	0.16	279	921
10	0.72	0.83	0.46	0.39	0.13	0.24	0.13	0.19	279	921

Notes: Graph shows gender wise work status of the survey respondent in the preceding week of the different survey rounds. Data are drawn at the survey respondent level; only matched respondents are used in the analyses. Farming includes self-employment in a household based crops cultivation, other farming tasks, or livestock activity. Business includes household based enterprise, for example, as a trader, shopkeeper, barber, dressmaker, carpenter or taxi driver. Wage/service includes work for a non-household member, fr example, an enterprise, company, the government or any other individual. Finally, any work is a combination of farming, business, and wage/service.

Table A.4: Effect of COVID-19 on the Probability of Employment (Respondent)

	(1)	(2)	(3)	(4)
	Any work	Farming	Business	Wage/services
COVID intensity	0.0273	-0.139*	0.108***	0.0631**
	(0.0278)	(0.0723)	(0.0395)	(0.0270)
Post	-0.0845***	-0.265***	-0.122***	-0.0182
	(0.0208)	(0.0474)	(0.0442)	(0.0167)
COVID intensity X Post	-0.0543*	0.220**	-0.256***	-0.0194
	(0.0282)	(0.0954)	(0.0745)	(0.0216)
Constant	1.036***	0.756***	0.447***	0.138***
	(0.0459)	(0.0565)	(0.0499)	(0.0363)
Observations	14,400	14,400	14,400	14,400
R-squared	0.123	0.073	0.094	0.137
Endline control mean	0.787	0.281	0.395	0.127

Notes: Data are drawn from the Nigerian Living Standard Survey and are at the respondent level. Post is a dummy variable taking a value of 1 for survey rounds taking place after the COVID-19 pandemic took place, and 0 otherwise. COVID intensity is a dummy variable taking a value of 1 if the household is from a State where COVID-19 infection rates were above the median value of the State-wise infection rate distribution during March 1st to April 30, 2020. Each regression controls for individual and survey round fixed effects. Each regression control for socio-economic controls which includes age, marital status, and education level of household head. Standard errors are clustered at the state level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.5: Effect of COVID-19 on the Probability of Employment by Gender (Respondent)

	(1) Any work	(2) Farming	(3) Business	(4) Wage/services
Female	-0.0382*	-0.160***	0.0873*	-0.0607
	(0.0221)	(0.0537)	(0.0484)	(0.0371)
Post	-0.103***	-0.208***	-0.204***	-0.0367***
	(0.0182)	(0.0510)	(0.0341)	(0.0133)
Female X Post	-0.0429*	0.236***	-0.208***	0.0378*
	(0.0253)	(0.0753)	(0.0668)	(0.0196)
Constant	1.084***	0.712***	0.504***	0.200***
	(0.0445)	(0.0570)	(0.0509)	(0.0434)
Observations	14,400	14,400	14,400	14,400
R-squared	0.125	0.070	0.079	0.134
Endline control mean	0.799	0.303	0.356	0.160

Notes: Data are drawn from the Nigerian Living Standard Survey and are at the respondent level. Post is a dummy variable taking a value of 1 for survey rounds taking place after the COVID-19 pandemic took place, and 0 otherwise. Female is a dummy variable taking a value of 1 if household head is female, and 0 otherwise. Each regression controls for individual and survey round fixed effects. Each regression control for socio-economic controls which includes age, marital status, and education level of household head. Standard errors are clustered at the state level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.6: COVID-19 intensity and the Probability of Employment by Gender (Respondent)

	(1) Any work	(2) Farming	(3) Business	(4) Wage/services
Female	-0.0294	-0.0563	0.0466	-0.0309
	(0.0314)	(0.0560)	(0.0453)	(0.0424)
Post	-0.0807***	-0.283***	-0.0940*	-0.0301
	(0.0231)	(0.0476)	(0.0468)	(0.0219)
Female X Post	-0.0156	0.0681	-0.105	0.0445
	(0.0318)	(0.0749)	(0.0662)	(0.0317)
COVID intensity	0.0315	-0.0938	0.0931**	0.0728**
	(0.0319)	(0.0720)	(0.0396)	(0.0344)
Female X COVID intensity	-0.0159	-0.253***	0.107	-0.0534
	(0.0427)	(0.0664)	(0.0659)	(0.0482)
Post X COVID intensity	-0.0413	0.142	-0.207***	-0.0129
	(0.0321)	(0.0881)	(0.0708)	(0.0261)
Female X Post COVID intensity	-0.0709	0.414***	-0.277***	-0.0181
	(0.0476)	(0.0901)	(0.0847)	(0.0411)
Constant	1.068***	0.760***	0.465***	0.152***
	(0.0523)	(0.0607)	(0.0524)	(0.0464)
Observations	14,400	14,400	14,400	14,400
R-squared	0.128	0.087	0.107	0.139
Endline control mean	0.799	0.303	0.356	0.160

Notes: Data are drawn from the Nigerian Living Standard Survey and are at the respondent level. Post is a dummy variable taking a value of 1 for survey rounds taking place after the COVID-19 pandemic took place, and 0 otherwise. Female is a dummy variable taking a value of 1 if household head is female, and 0 otherwise. COVID intensity is a dummy variable taking a value of 1 if the household is from a State where COVID-19 infection rates were above the median value of the State-wise infection rate distribution during March 1st to April 30, 2020. Each regression controls for individual and survey round fixed effects. Each regression control for socio-economic controls which includes age, marital status, and education level of household head. Standard errors are clustered at the state level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Covid Economics 82, 23 June 2021: 70-98

The importance of being earners: Modelling the implications of changes to welfare contributions on macroeconomic recovery[1]

Max A. Mosley[2]

Date submitted: 16 June 2021; Date accepted: 20 June 2021

This paper demonstrates how changes to welfare generosity during recessions induces a greater than usual economic response. This is predicated on the assumption that welfare recipients are likely to be liquidity-constrained and therefore highly responsive to a change in temporary income. This would result in two conclusions, (i) the effects of fiscal stimulus can be maximised when channelled through welfare and (ii) fiscal consolidation from these programs will have a strong contractionary effect on domestic output. Using tax-benefit microsimulation model UKMOD, we find 71% of means-tested welfare recipients are liquidity-constrained. We use this finding to calibrate an open-economy New Keynesian macroeconomic model to therefore illustrate the economic implications of positive changes to the program's generosity, finding an impact fiscal multiplier of 1.5. For cuts to contributions, we find a negative multiplier of 1.8, implying past cuts to welfare had a sizeable contractionary effect on macroeconomic recovery.

Covid Economics 82, 23 June 2021: 99-143

1 I am particularly grateful for the guidance, support and engagement Professor Andr s Velasco has given this work. I would also like to thank Dr Evan Tanner, Dr. Iva Tavessa, Professor Kitty Stewart, Sir Julian Le-Grand, Charles Morris and Sam Anderson for their individual contributions to the ambitions of this paper. The results presented here are in part based on UKMOD version A2.50+. UKMOD is maintained, developed and managed by the Centre for Microsimulation and Policy Analysis (CeMPA) at the University of Essex. The process of extending and updating UKMOD is financially supported by the Nuffield Foundation. The results and their interpretation are the author's responsibility.

2 MPA Candidate in Economic Policy at the LSE.

1. Introduction

Following major contractions in output, governments can stimulate economic activity by purchasing infrastructure or cutting taxes/transferring cash to defined households. The need for such interventions has been made more necessary while monetary policy – which has the power to pull back on or even fully offset any expansionary effect of fiscal stimulus – remains constrained at the zero-lower-bound; this has resulted in fiscal policy taking a renewed frontline role as a stability mechanism (Shoag, 2013).

The efficacy of cash-transfers – which are often the most appropriate method of stimulus delivery due to the comparatively short implementation time – is dependent on recipients choosing to spend the windfall. However, the Barro-Ramsey model in standard consumption theory predicts that temporary income variations will not induce a consumption response as households will save any temporary windfall in anticipation of a future tax rise to pay for it (Barro, 1974). This suggests such households have a marginal propensity to consume (MPC) of 0. If stimulus is not spent and instead saved in its entirety, the ratio of the output increase to stimulus, known as the fiscal multiplier, will be at or close to 0, rendering the intervention ineffective.

This simple model assumes that all households have equal access to alternative sources of cash (savings) or debt (credit markets) to act as a buffer to any income shock to allow the household to finance a permanent level of consumption (Canbary & Grant, 2019). However, a large amount of empirical literature, starting with Hall (1978), has consistently found 20% of households do not adhere to this permanent income hypothesis because they have little savings and/or are excluded from credit markets (hereafter referred to as liquidity-constrained). This inability to draw on alternative sources of liquidity shortens the horizon for financial planning (Campbell & Hercowitz, 2019), resulting in this subset of households being therefore highly sensitive to a change in temporary income (Jappelli, et al., 1998; Jappelli & Pistaferri, 2014; Parker, Souleles and Johnson, 2006; Johnson, et al., 2006). As such, papers that model the fiscal multiplier only for liquidity-constrained households often find strong responses, with Kenichi Tamegawa (2012) concluding:

Covid Economics 82, 23 June 2021: 99-143

"The maximum value of the multiplier is obtained when the share of liquidity-constrained households is close to unity" (Tamegawa, 2012)

If liquidity-constrained households are the strongest – and arguably sole – demand-side channel for stimulus, can governments maximise its multiplier by making it available only for these defined households? Historic cash-transfers have, to the best of our knowledge, never been made available to just households with low savings/credit market access. This is likely for two reasons: firstly, it would take a large administrative effort to identify households who meet this criterion, involving a lengthy and dangerous delay while governments audit each household's total financial assets. Secondly, it would likely be too politically difficult to justify transferring stimulus to these households exclusively as to the general population this could appear to be a somewhat arbitrary criterion for stimulus checks.

Consequently, if the onus for stabilising short-term outcomes has fallen on cash-based fiscal stimulus, but this can only influence economic activity when liquidity-constrained households gain, the policy is at best inefficient if we currently have no realistic way to target them specifically. Many therefore argue fiscal stimulus to be 'too circumscribed' (Cochrane, 2010) if it can only influence a small subset of the population. Cochrane's challenge to any proponent of fiscal stimulus is that they must either disprove the claim that the majority of households consume from their permanent income or find a way for stimulus to better target these liquidity-constrained households.

This paper assesses the role of existing welfare programs in meeting this latter challenge, as means-tested social assistance, by definition, is only available for households with low savings. For instance, the United Kingdom's (UK) Universal Credit scheme's strict criteria means that recipients can only claim state assistance if a household's total savings are less than £16,000 (DWP, 2021). Therefore, these programs appear naturally designed to benefit liquidity-constrained households. If this is the case, the following conclusions would result.

Firstly, fiscal multipliers can be maximised when stimulus is channelled through these programs as they would present the most effective way to transfer cash directly to the

Covid Economics 82, 23 June 2021: 99-143

households who are liable to spend it. This presents little administrative challenge, as raising the levels of existing structures can be enacted quickly; the £20 boost to Universal Credit was enacted a few weeks following COVID-19 restrictions (HM Revenue & Customs, 2020). This would also likely be politically feasible, as public support for raising welfare levels during recessions is usually high, with 74% of the public being in favour of the aforementioned boost to Universal Credit (Ipsos MORI, January 2021).

Secondly, fiscal consolidation from cuts to welfare will induce a stronger contraction in domestic consumption and thus output. The Bank of England provide one of the only estimates of MPCs from both increase and falls in income; they find consistently higher estimates from the latter than the former (Bunn, et al., 2017). If liquidity-constrained households are not only the most responsive to a positive change in temporary income but are even more responsive to negative shocks, this implies that consolidating welfare spending would induce a strong contractionary effect on domestic consumption.

This is not the first paper to assess the role of liquidity constraints in strengthening fiscal multipliers but is, to the best of our best knowledge, one of the first in assessing the role of welfare programs in achieving this goal. We believe the study of fiscal multipliers out of welfare programs has only ever been studied once before by Gechert et al (2021) for Germany, who note a similar dismay at the lack of academic attention given to the question. They opt for the popular (s)VAR strategy to estimate the multiplier of exogenous shocks to welfare implementation, finding consistent multipliers of 1.1 as a result of the strong representation of liquidity-constrained households in the program. The AARP similarly studied the general question of how welfare is connected to the domestic economy by using an 'off-the-shelf' impact assessment model IMPLAN to measure this, finding it supported $1.4 trillion in output in one year (Koenig & Myles, 2013). Though compelling, the approach fundamentally lacked any econometric detail, asking the reader to focus solely on the outcome and forego any consideration for how it was arrived at.

Covid Economics 82, 23 June 2021: 99-143

The first contribution this paper seeks to make arises by centring its analysis on the United Kingdom, where existing welfare generosity ranks low compared to other European nations. Specifically, the UK's replacement rate, that is the proportion of average income replaced by unemployment benefits,[1] ranks the lowest on a range of measures (Spinnewijn, 2020). This is important, as the reader could agree that welfare programs provide a strong avenue for stimulus but claim that this already happens following a recession when more people become eligible for welfare, known as an automatic fiscal stabiliser. But the ability for the UK's automatic stabiliser to enact the above is weak if its welfare programs are already meagre, meaning the UK government cannot rely on its existing programs to stimulate demand without an additional stimulus boost. As a result, it is common that countries with low automatic stabilisers (such as the UK) enacting higher levels of stimulus during economic crises (Dolls, et al., 2012) and vice versa.

Our paper first confirms the fundamental assumption that welfare programs already target this strong demand side channel by determining how many liquidity-constrained households benefit from the program compared to tax-cuts as an alternative. We do this by using tax-benefit microsimulation model UKMOD which can simulate the distributional consequences from changes to both welfare and tax levels using data from the 2018 Family Resources Survey (FRS). We simulate the effects of changes to welfare policies and tax-rates and compare the number of liquidity-constrained households that gain. As this is a static model, it can only show the 'morning after' effects of a policy or policy reform and cannot initially solve for core macroeconomic outcomes such as the relationship between the program and demand stimulation. This paper therefore takes a novel approach to UKMOD, by using it to test core assumptions we can then use to build an accurate macroeconomic model.

The macroeconomic model we opt for is an open-economy New Keynesian extension of an IS/LM setting created by Tanner (2017). We use this model to solve for core macro variables such as the output gap and thus draw inferences about the multiplier effect

[1] Spinnewijn measures at the start of an unemployment spell for a representative 35-year-old worker with an employed partner and one child earning the respective countries average salary before unemployment spell.

Covid Economics 82, 23 June 2021: 99-143

from different fiscal policy designs. As this model is simpler and relies on a number of exogenous parameters, we provide transparent robustness checks that calibrate the model so that its outputs are consistent with historical outcomes. We assess the size of the multiplier from positive and negative changes to welfare contributions on core macroeconomic outcomes, consumption/investment/net exports etc.

Typically, papers of this nature would attempt to estimate effects using either a quasi/natural statistical experiment or use comprehensive DSGE/(s)VAR models. Regarding the former, changes to welfare contributions happen at either micro-level (such as the regional roll-out of a new welfare program) where there is insufficient micro-data to determine the causal effect from, or the macro-level (country-wide) where it is not possible to disentangle the effect of the welfare reform from other economic factors. Papers instead often opt for the latter set of sophisticated economic models which can estimate comprehensive, dynamic economic outcomes following hypothetical policy shocks. However, such complex models are naturally computationally intensive, making them particularly inaccessible to even seasoned economists (Krugman, 2000). This has brought them into sharp criticism by high profile economists, including Blanchard (2009) and Romer (2016) for their inability to communicate salient economic policy to policy makers. These models are therefore better suited in providing evaluations of economic outcomes for more academic audiences.

The simpler static model employed in this paper can aid more transparent communication of the key macroeconomic relationships and outcomes to a potentially non-technical audience. This approach aims for something of a middle-ground between the two methods above, to test for and demonstrate the intuition of this paper. However, what we gain in transparency we lose in economic precision, so this paper can be seen as an illustration of this position regarding the role of welfare programs in fiscal policy. This gives future papers in this area with more robust models a benchmark to compare results to.

This paper takes a novel approach to optimising fiscal stimulus, assessing the role of existing welfare programs by using transparent and intuitive econometric methods. We

Covid Economics 82, 23 June 2021: 99-143

Covid Economics 82, 23 June 2021: 99-143

also extend this position to estimate the contractionary effect of cuts to these programs, thus providing a comprehensive account for how changes to the generosity of welfare contributions can influence economic recovery. The paper is organised as follows. A conceptual framework in section 2 will outline key literature on fiscal stimulus and liquidity-constraints. Section 3 will outline the methodology for both the microsimulation technique and the key features of the macroeconomic model. Our findings will then be split the outputs from the microsimulation and the macroeconomic model in section 4 and 5. The implications from both sets of findings are considered in the discussion in section 6.

2. Conceptual Framework

2.1 Review of Literature and Debates Over Fiscal Policy

How effective fiscal policy is in influencing the economic outcomes has been long debated by economists. Investigations of historic fiscal multipliers over the post-war period have taken broadly two forms of inquiry; first, papers that track the observed economic effects of exogenous build ups of post-war military spending as a natural experiment; finding multipliers ranging from 0.6-1.6 (Edelberg et al ,1998; Hall, 2009; Ramey, 2009; Nakamura & Steinsson, 2011). The second kind utilises structural vector autoregressions (SVAR) to empirically test for past multipliers and its determinants, finding multipliers from 1-1.5 (Blanchard and Perotti, 2002; Ramey, 2011; Gechert, 2021). More recently, the 2008 US stimulus package was prominently stated to have a multiplier of 1.6 by the Chair of Council of Economic Advisers to President Obama (as cited in Ilzetzki, et al., 2013). This drew sharp criticism from Robert Barro, who argued the output multipliers are near 0 as the gains from government purchases are partially or fully offset by the negative impacts they have on private investment (Barro, 2009). Barro later calculated that the extra $600bn in this stimulus spending came at the cost of $900bn fall in private investment (Barro, 2010), implying a multiplier of just 0.6.

How do we reconcile these competing views? It could be that the economic environment today is no longer as hospitable to fiscal interventions as it was in the post-war period. Ilzetki et al (2013) provide evidence for this, by identifying the key characteristics that determine the size of these historic spending multipliers by employing the same SVAR strategy as Blanchard & Perotti (2002). One notable feature they find of multipliers is that they are strongest in low-debt (<60% debt to GDP) countries. The far less debt in the post-war period compared to Barro's time-frame perhaps provides an answer as to why Blanchard and Perotti find a higher multiplier result. The evolving economic environment, from post-war low- to high-debt economies, could be argued to have initially shifted the consensus on the efficacy of fiscal policy away from large multipliers to more conservative estimates.

But although worldwide debt-GDP has remained high, nominal interest rates have now been persistently constrained at the zero-lower-bound (ZLB) since the great recession. This one factor alone has significant implications for the future of fiscal interventions, as with weakened monetary policy, fiscal multipliers have been shown to be significantly higher (Christiano, et al., 2010; Erceg & Lindé, 2014), with Hall estimating multipliers of 1.7 (Hall, 2009). In normal times, monetary policy leans on fiscal expansions by raising interest rates to increase the cost of borrowing for firms, and returns to saving for households, which reduces private investment and consumption, thus reducing the size of the multiplier. But when monetary policy is already at its minimum value, the proverbial 'brakes' are already released on the economy, creating for a highly responsive environment to fiscal policy, and therefore higher multipliers (ibid). The question becomes which form – government purchases, tax-cuts or cash-transfers – is most effective in stimulating economic activity? There are two considerations in determining efficacy; the time it takes to enact the stimulus; and the amount of windfall spent by households.

Regarding the former, speed of implementation is vital for recovery strategies, as without a strong monetary response, economies will be in near freefall until fiscal policy can be executed. The longer the implementation-lag, the deeper the recession (Tsurugaa & Wake, 2019). This issue is not exclusive to government purchases, which are often infrastructure based and therefore slow, as tax-cuts can similarly only be delivered to households according to the natural tax schedule (Romer & Romer, 2010). For short-term recovery, governments instead opt for cash-transfers at the onset of the recession to avoid these implementation lags. For instance, the United Kingdom was able to enact its furlough program in 2020 three days before lockdown even began.

The efficacy of cash-transfers will depend on the amount of the temporary transfer that is spent by households, measured by the marginal propensity to consume (MPC). Economists have long been sceptical that households would ever spend a temporary gain (implying an MPC of 0), as the majority of households will only consume out of their permanent, as opposed to temporary, income and therefore save the entirety of the windfall in anticipation of future tax rises to pay for the stimulus (Barro, 1974; Cochrane, 2010). Although this has been found to apply to well over the majority of households (Hall & Mishkin, 1982; Canbary & Grant, 2019), households with low-savings and little access to credit markets are in fact highly responsive to both positive and negative temporary income changes (Johnson, et al., 2006). For these households, studies have estimated MPCs as high as 0.9[2] (Canbary & Grant, 2019) as they cannot smooth consumption out over the life course.

2.2 Defining Liquidity Constraints

There is some variation in how previous studies have formally defined a liquidity-constrained household, given the fact 'low savings' is an ambiguous term. There are four compelling sets of criteria that attempt to isolate households from sources of plausible earnings: (i) savings, (ii) market earnings, (iii) home-owners or (iv) credit markets. The first is captured by the Zeldes definition, which classifies liquidity constraints as households with total wealth of less than two months disposable income. Although this neatly captures a lack of savings relative to a household's given earnings, measurement of household wealth is often prone to error and datasets often do not collect it for this reason (Jappelli, et al., 1998). Further, this definition only works if the relationship between wealth and liquidity constraints is perfectly monotonic (Dolls, et al., 2012).

Runkle (1991) therefore focuses on the second and third sources by considering all unemployed households without a mortgage as liquidity-constrained. The clear logic

[2] Meaning for these households, 90% of the income gain will be spent in the domestic economy

behind this approach is that unemployed households (where there is no adult working) with no income and no ability to liquify the capital stored up in their home will have little opportunity to smooth out the temporary income shock. The fourth is perhaps the most difficult to obtain data for, as credit market statistics will be held only by private stakeholders. It is therefore common to use survey data that directly asks households about their access to credit, as Jappelli et al (1998) and Dolls et al (2012) do, such as with the FCA financial lives survey which asks participants if they have had a rejected credit application (FCA, 11 February 2021, p. 123).

Due to data and methodological constraints explained below, we opt for a combination of the second and third measure. As we will be using tax-benefit model UKMOD to determine which forms of fiscal stimulus target liquidity-constrained households the most, we are therefore limited by the variables available in Financial Resources Survey (FRS) dataset the model relies on, meaning we cannot at this stage use the FCA dataset. Unfortunately, there is little data on savings/wealth, meaning we cannot take the first approach in its entirety. Instead, we build off the second approach, identifying all households who do not own their own home and with no working adults[3] to be liquidity-constrained. Lastly, we include a fifth source of income; (v) the household having no investment income.

2.3 The MPC for Liquidity-constrained Households

Many papers that attempt to identify MPCs for households from transitionary gains often differentiate between representative and liquidity-constrained households for this reason, as the consumption response for a household with low savings and without credit market access will be higher than the population. A summary of this literature is presented in Table 1, which shows that MPCs are consistently found to be higher for liquidity-constrained households than typical households under range of scenarios and country-settings. For each study, we see far higher MPCs when looking at just liquidity-constrained households than at the overall population.

[3] We drop this unemployment requirement when looking at tax-cuts, explained below

Table 1: Literature Estimates of MPCs

Author(s)	Context/Sample	MPC Estimates		Notes
		Overall	Liquidity-constrained	
Agarwal and Qian (2014)	2011 Growth dividend (Singapore)	0.8	0.5-0.75	Estimate is at both announcement & dismemberment
Johnson, Parker and Souleles (2006)	2001 US Income Tax Rebates	0.2-0.4	Larger	First of many papers that uses random timing of stimulus-based welfare number
Johnson, Parker Souleles and McClelland (2013)	2008 US Stimulus Payment	0.5-0.9	Larger	Same method as above
Tullio Jappelli & Luigi Pistaferri (2014)	2010 Italian Dataset	0.48	0.7	Low 'cash-on-hand' households exhibit larger MPCs
Zara Canbary and Charles Grant (2019)	1986-2010 UK FRS Dataset	0.5-0.94	0.75-0.94 (higher following recessions)	Find only 50% of households consume from permanent income
Fisher et al (2019)	1999-2013 US PSID Dataset	0.2-0.6	Larger	MPC tapers off to 0 after the 3rd wealth quintile
Tal Gross, Matthew Notowidigdo, and Jialan Wang. (2016)	US Consumer Credit Panel (CCP)	-	0.37 (20-30% higher during great recession)	Measure effects of bankruptcy flag removal on consumption
Crossley et al (2021)	Survey over COVID-19	0.11	-	Do not test for liquidity-constrained households specifically

Despite the fact these estimates are found from variance in study type, each consistently shows that the MPC is highest for liquidity-constrained consumers. This is likely because the basic intuition is the same, households with low access to alternative sources of cash will be responsive to temporary income changes. For our analysis, we take Canbary and Grant's estimates as true representations of the MPC for liquidity-constrained households as their estimates are for the UK (thus eliminating the effects of any country specific factors) and is estimated from the dataset we use in UKMOD. We then apply this MPC to the percentage of liquidity-constrained households within welfare programs identified in our microsimulation exercise. For these liquidity-constrained households we can assume with some certainty a change in temporary income will induce a strong change in consumption as they do not have alternative income sources to draw on.

Existing literature has identified the need for cash-based fiscal stimulus and specified the households it needs to target, but there is a gap in understanding how best to achieve this. Therefore, this paper explores the role of welfare policies in meeting this challenge and will evaluate if existing programs already benefit liquidity-constrained households. If so, we will be able to illustrate the economic consequences of positive and negative changes to welfare using a macroeconomic model.

3. Methodology

For studies of this nature, there are two possible methodological candidates. The first is to test for economic outcomes following real-world changes in welfare contributions by determining their causal effect using quasi/natural experiments. However, this has not been possible due to data constraints, as the changes we can track are on a micro scale whereas the available data on economic indicators (such as consumption levels) are aggregated. Further, this approach does not give us a plausible avenue to explain why we observe a given outcome. This paper is based off existing economic intuition about the role of liquidity-constrained households in strengthening the effects of fiscal stimulus, but quasi/natural experiments do not give us the opportunity to determine if it is this that is driving our results or if it is being driven by some other factor.

Instead, studies of this sort opt for the alternative class of methodologies is through the use of sophisticated economic models such as with a DSGE of (s)VAR framework. Though powerful, these computationally intensive methods struggle to communicate results to a non-statistical audience (Krugman, 2000) and have therefore been criticised for their lack transparency (Romer, 2016).

Our methodological approach has been chosen as something of a middle-ground between these two approaches, that is built off real-world observations and estimates results using a less intensive New Keynesian extension of an IS/LM macroeconomic model. Therefore, our methodology is split into two approaches. The paper first tests for

Covid Economics 82, 23 June 2021: 99-143

the fundamental assumption that welfare programs target liquidity-constrained households through the use of microsimulation model UKMOD. We then estimate the fiscal multiplier effects of a hypothetical change in the levels of contributions using this macroeconomic model.

3.1 Microsimulation through UKMOD

We first confirm the fundamental assumption that existing welfare programs target liquidity-constrained households using the tax-benefit microsimulation model UKMOD. This model is built from the 2018 Financial Resources Survey (FRS) which provides figures on the personal and financial characteristics of the population, and welfare recipients specifically. This official dataset provided by the Office for National Statistics (ONS) is a continuous survey of UK households, comparable to EU-SLIC, which provides statistics on income sources and general characteristics including home ownership.

The UKMOD microsimulation model opens up opportunities to simulate tax-benefit changes and assess the distributional consequences. Specifically, we can simulate an increase in benefit levels and determine what proportion of those who gain are liquidity-constrained. We therefore code 1 for those who see an income change and 0 otherwise. We apply this analysis to each type of UK benefit to determine the presence of any heterogeneity across programs. It is likely that means-tested benefits are better able to target liquidity-constrained households than non-means tested benefits, as the former is designed to specifically target financially precarious households whereas the eligibility for the latter is not necessarily savings/credit related (e.g., child benefit). This UKMOD model focuses analysis on taxes and benefits applied to the whole of the UK and assumes full benefit take-up and tax-compliance at the household unit level. As this is a controlled microsimulation experiment, we do not have to worry about endogeneity issues, as we are able to isolate the effects of the given changes to taxes and/or benefits. Improvements to this approach are specified in section 6, specifically how new datasets can be added and more analysis across different time periods would improve the overall precision of our estimates and their generalisability.

Covid Economics 82, 23 June 2021: 99-143

We can compare these results to cuts in tax-rates to see how many liquidity-constrained household's gain. As it is common for stimulus through tax-cuts to be targeted at 'lower' income households, such as the 2001 US tax-rebate (Johnson, et al., 2006), we therefore simulate this tax-cut on the lowest those tax-bands; again coding 1 for those who see an income change and 0 otherwise. As mentioned above, there are a number of ways to define 'liquidity-constrained' households and we include 'unemployment' as a key feature. However, for tax cuts this would result in no liquidity-constrained households gaining as they would not be earning market income under this strict definition. Therefore, for the tax-cut stimulus simulation we drop the unemployment criteria (keeping the no investment and/or non-homeowner measure).

This can provide a robust account of how different designs of stimulus can target more or less liquidity-constrained households which is presented below. This can therefore help us identify what is the appropriate average MPC for welfare recipients or those who benefit from a tax-cut.

3.2 Constructing a Macroeconomic Model

If welfare programs do target liquidity-constrained households, we can illustrate the implications of stimulus through these strong demand-side welfare programs by calibrating a macroeconomic model. The model we use is detailed in Appendix A and specified in full in Tanner (2017), but here we outline how the IS curve, interest rates and output gap are calculated. This allows us to produce multiplier estimates of fiscal shocks measured as a percentage of potential output. First, we substitute the rescaled equations in Appendix A for consumption, investments and net exports into a New Keynesian GDP identity, including a measure for government purchases gp_t:

$$Y_t = Y^P * [1 + \left(1 - \sigma_{cyc}\right)\{[(1 - \tau)(gap_t)] - tp_t\} + \varphi_{Ir}(r_t - \overline{\mathrm{r}}) + gp_t]$$

φ_{Ir} is a response parameter scaled to potential output, so that $\varphi_{Ir} = \alpha_r / Y^P$. We then subtract and divide both sides by potential output to solve for the output gap IS curve:

$$gap_t = \frac{\varphi_{Ir}(r_t - \bar{r}) + gp_t - (1 - \sigma_{cyc})tp_t}{\tilde{\sigma}_{cyc}}$$

Note, this by construction takes the form of the traditional Keynesian multiplier (1/savings), as $\tilde{\sigma}_{cyc} = 1 - (1 - \tilde{\ }\sigma_{cyc})(1 - \tau) = 1/\tilde{\sigma}_{cyc}$. We can differentiate between fiscal policy designs by firstly including two different parameters for fiscal intervention, gp_t government purchases and tax policy tp_t (cash transfers therefore the inverse of tp_t. We then solve to include net exports which is captured by the terms-of-trade parameter TT_t.

$$gap_t = \frac{(\varphi_{Ir} - \eta_{nx})(r_t - \bar{r}) + fp_t - \tilde{\eta}_{nx}\ln(TT_t)}{\{1 - [(1 - \sigma_{cyc})(1 - \tau) + im_{cyc}]\}}$$

Both measures of fiscal policy are now summarised into one identity $fp_t = gp_t - (1 - \sigma_{cyc})tp_t$. Here, our response parameters summarise the above by $\eta_{nx} = \eta_x - \eta_{im}$ and $\tilde{\eta}_{nx} = \eta_x(1 - \nu) + \eta_{im}\nu$. We can think of TT_t as foreign demand, as improvements in trade terms improve the IS curve will shift to the right. We solve for equilibrium output by first flipping the above to create an expression for the real interest rate:

$$r_t = r^{NAT}\,\frac{\{1 - [(1 - \sigma_{cyc})(1 - \tau) + im_{cyc}]\}gap_t - [fp_t - \tilde{\eta}_{nx}\ln(TT_t)]}{\varphi_{Ir} - \eta_{nx}}$$

We then solve for the equilibrium output gap by combining the above equation with the equilibrium of real interest rates:

$$gap_t^{eq} = \frac{\beta_\pi(\pi^e - \pi^t) + \beta_{ss}(ss_t) + efp_t\left[\beta_{efp} + \frac{\eta_{nx}}{(\varphi_{Ir} - \eta_{nx})}\right] + \frac{[fp_t - \tilde{\eta}_{nx}\ln(TT_t)]}{(\varphi_{Ir} - \eta_{nx})} + r^{disc}}{\left[\frac{1 - [(1 - \sigma_{cyc})(1 - \tau) + im_{cyc}]}{(\varphi_{Ir} - \eta_{nx})} - \beta_{gap}\right]}$$

We simulate the effects of welfare expansions by imputing a one-off tax policy tp_t of -1% and solving the for the effects on output (consumption, investment etc.) by comparing the percentage change from period 1 to period 2 (the latter with the policy shock). We input this into a demand shock component which eventually interacts with the MPC $(1 - \sigma_{cyc})$. This builds this intuition that MPC size is what leverages the size of the economic response. The inclusion of the MPC estimate allows us to model effects of any scenario based on its distributional effects to different households. This can show the implications for fiscal policy to both Ricardian equivalence/permanent income households with an MPC of 0 or liquidity-constrained households with an MPC>0 defined using Table 1. Further, this model can show how fiscal policy can appreciate or depreciate the real exchange rate depending on the relative strength of the monetary response in a more transparent and intuitive way than complex DSGE methods. This is of key importance for this paper, as it can show in a credible way how different monetary conditions (ZLB) can strongly influence the effectiveness of fiscal policy.

We can therefore compare the effects of the expansion in a number of scenarios which are (i) during a minor recession where monetary policy has room to respond to both the recession and expansion, (ii) during a major recession at the zero-lower-bound with a capital outflow scenario. We then replicate this latter analysis by decreasing welfare contributions.

Of course, this model does have a number of notable constraints that limit what can and cannot be inferred from its estimates. Firstly, this is only a static model in the same form as UKMOD, meaning we cannot forecast into the future what the outcome will be after multiple rounds of spending. Many papers that estimate fiscal multipliers do this (see Blanchard & Perotti, 2002; Ilzetki et al, 2013) to see how long it takes for the effect to equalise; we are unable to make such analysis from this model. Secondly, such New Keynesian models have multiple exogenous variables which are independent from the policy change. Notably, our nominal exchange rate S_t (see Appendix A) is exogenously defined, meaning it is not connected to the given economic conditions so an appreciation in the current account following stimulus will not result in capital inflows. We therefore incorporate this with a forcing variable by decreasing external financial pressure by 0.1% to induce a capital inflow scenario, consistent with the set-up Tanner (2017) performs. For simplicity we assume this inflow is linear across MPC estimates, when in reality it will be dependent on the output response. Although this may help us avoid overestimating some results, the approach is imprecise in nature and is less convincing than a model that is able to do this naturally.

The 'Lucas Critique' refers in part to this problem of believing elements of structural equations to be exogenous, as even aspects of consumption are never truly independent from government policy (Sargent, 1987). As such papers opt for the more sophisticated models of SVAR and DSGE models mentioned earlier which are far better able to overcome these limitations.

Therefore, this model should be taken as an economic illustration of the above argument rather than a direct forecast for the UK. It would be a worthwhile exercise to cross-check this model with one of these other approaches, as Pappa et al (2015) do when assessing the impact of tax avoidance on fiscal consolidation. Such further research opportunities are addressed in section 6.

3.3 Robustness Checks

Because of these fundamental limitations in the precision and interpretability of our model, we believe it necessary to disclose its relative power by cross-checking its outputs from historical events to what they were in reality. We do this on the key variables that exert the strongest influence over our model, real and nominal interest rates along with inflation shown in Figure 1. We take historical data on output gaps calculated by the Office for Budget Responsibility (OBR, 2020) and past estimations of the natural rate of interest by Goldby et al (2015) and input this into the model to allow it to forecast what the outcomes would have been in the past and compare them to reality.

This exercise can also be helpful in choosing the values of certain exogenous parameters, such as those that make up the central bank's Taylor Rule or fundamental features of an economy such as the elasticity of short run aggregate supply. We therefore choose a number of parameters to minimise the distance between historic outputs and the predictions of our model. In doing so we obtain the following parameters in Table 2, taken either from existing papers or calibrated ourselves.

Covid Economics 82, 23 June 2021: 99-143

Table 2: List of Exogenous Parameters

Parameter	Description	Source	Value
π^e	Inflation expectation	Assumption	0.02
π^t	Inflation target	Assumption	0.02
β_π	Inflation weight	Author's calibration	2
β_{gap}	Output gap weight	Author's calibration	1
β_{ss}	Supply-shock weight	Author's calibration	1
η_{SRAS}	Short run elasticity of aggregate supply	Author's calibration	2
r^{NAT}	Natural rate of interest[4]	Evans (2020)	0.015
im_{cyc}	Short-run propensity to import	Author's calculation[5]	0.126
tp_t	Tax policy (one-off)	Policy shock	-0.01
r^{disc}	Deviation from Taylor-Rule[6]	Assumption	0
efp_t	External financial pressures	Assumption	-0.01
θ	Transmission of external shock to exchange rate	Tanner (2017)	0.10
σ_{cyc}	Marginal propensity to consume (MPC)	Canbary and Grant (2019)	0.857[7]
ν	Importance of imports	Tanner (2017)	0.03
η_x	Response function to export prices	Tanner (2017)	0.72
η_{im}	Response function to import prices	Tanner (2017)	-0.72
τ	Tax share[8]	IFS (2019)	0.25

For nominal interest rates in Figure 1 panel a, we are able to track pre and post zero-lower-bound levels well. Our model naturally 'recommends' highly negative interest rates, as it does not consider zero to being a limiting factor like a central bank will would. As such, in our future estimates below we set up the model to stop itself at 0 if the output from interest rate policy is negative to avoid creating highly negative interest rates.

If we simply subtract inflation from this predicted output (0 if negative), we obtain the following measure of 'real interest rates' in panel b, which if we compare to the same from actual outputs (real inflation subtracted from nominal rates) we see a high degree of similarity. We could compare this to actual real interest figures, such as those done by

[4] For robustness checks we use yearly estimates rather than this long-run value
[5] See Appendix B
[6] Although we consider this 0, we use this for force adjustments where necessary to keep interest rates ≥0
[7] Calibrated as an average of their estimate range for liquidity-constrained households
[8] Calibrated to reflect low-income households

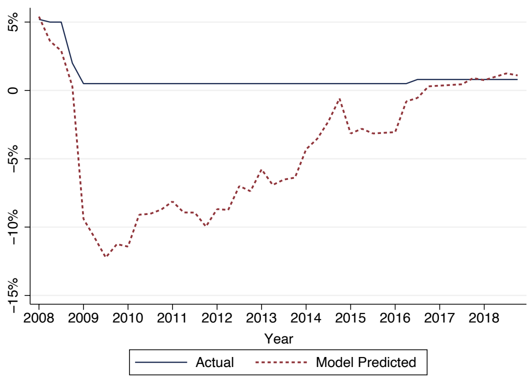
Covid Economics 82, 23 June 2021: 99-143

the World Bank, however their inflation deflators are not the same as ours so the outputs would not be interpretable.

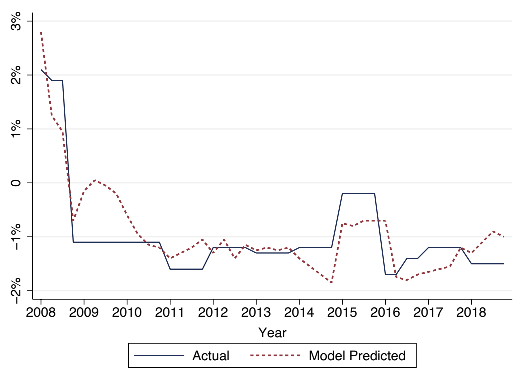
We simulate the effect of the 2015/16 oil shock by imputing a supply shock (ss) of 1%, as this episode had a substantial effect on inflation (Bank of England, 2016). As such, although our inflation in panel c is able to track real inflation fairly well, supply shocks must be simulated manually as the model cannot predict this itself otherwise it would have not noticed the 2015/16 oil price shock as this would not have been that well reflected in the output gap.

Figure 1: Robustness Checks

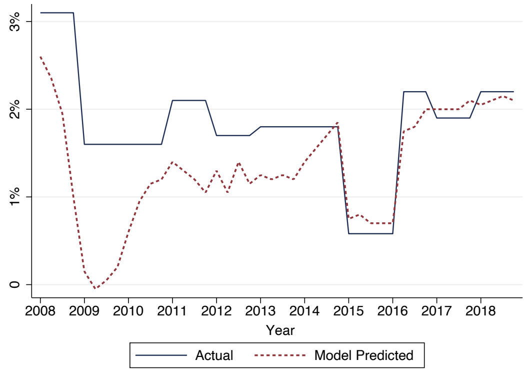
(a) Nominal Interest Rate



(b) Real Interest Rate



(c) Inflation



Overall, although our model is constrained in a number of areas, it is able to match historic outputs fairly well, at times with some adjustments. We are confident that this gives us a credible basis to illustrate the effects of exogenous changes to fiscal policy through welfare.

4. Findings: Microsimulation

The microsimulation exercise using UKMOD allowed us to first estimate that 30% of all households are liquidity-constrained using the criteria mentioned in Section 2, which is similar to Hall's (1978) 20% estimate. This at first confirms the concern that fiscal stimulus targeted at the broad population will be inefficient at targeting liquidity-constrained households.

4.1 Liquidity-constrained households by benefit category

When we simulate the effects of a change in welfare contributions, we see the following distributional consequences for liquidity-constrained households[9]. From broad welfare programs, we find in total 58.4% recipients are liquidity-constrained. But when we start to look within the different welfare programs that make up this finding in Figure 2, there is some important variations. First is the difference between means-tested and non-means-tested programs.

We take the difference between child tax credits and child benefits as an example of this, as both are similar in design but only the former is means-tested. For the non-means-tested program, only 39% of recipients can be classified as liquidity-constrained compared to 63% of recipients from the means-tested equivalent. This is consistent with the intuition of this paper that the reason welfare programs can target liquidity-constrained households is because the criteria to be eligible for means-tested welfare is very similar to what we would consider a household to be liquidity-constrained (e.g., having low levels of savings). When we look at the rest of the means-tested programs, we see a consistently high proportion of recipients being liquidity-constrained. Notably, we can infer that 91% of households impacted by the Universal Credit cap are liquidity-

[9] These are the same for both positive and negative changes to contributions

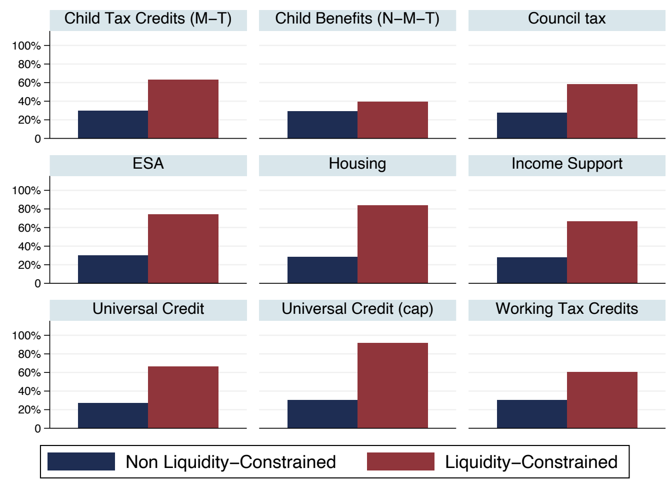
Covid Economics 82, 23 June 2021: 99-143

constrained. Overall, we find 71% of means-tested welfare recipients are liquidity-constrained, but just 27% are in non-means-tested programs; similar to the wider population. This intuitive, as without strict eligibility criterions in non-means-tested programs the demographic make-up of recipients will more broadly reflect the population.

4.2 Liquidity-constrained households by Tax Cut

We then repeat this exercise by simulating the effects of a tax-cut to lower-income households to provide some contextual clarity to the above finding. Specifically, we are interested in determining if the above implication is limited to just welfare programs, or if tax-cuts also have this ability to benefit liquidity-constrained households.

Figure 2: Liquidity-Constrained Recipients by Welfare Program



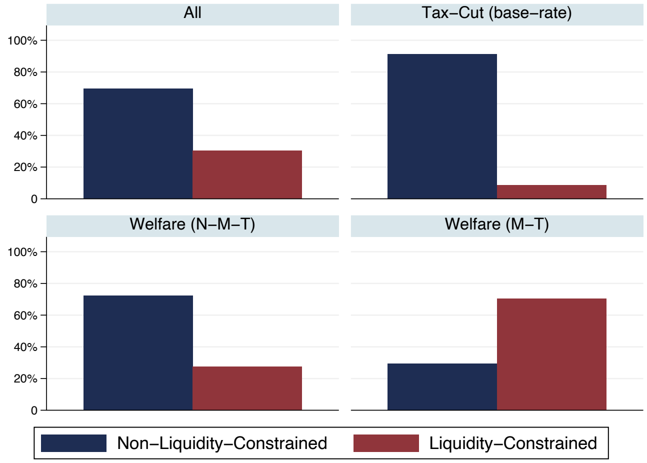
Source: Authors Calculation through UKMOD

We use UKMOD to simulate a tax-cut through from a 1% reduction in the liability within lowest basic-rate (£12,571-£50,270) tax band. We again code households who see a rise in disposable income 1 and solve for liquidity-constrained and non-liquidity-constrained households. We only simulate a tax-cut for the bottom band to make for a plausible comparison for stimulus through welfare programs. Our analysis in Figure 3 shows tax-cuts, even on the lowest income band, are particularly inefficient in benefiting

Covid Economics 82, 23 June 2021: 99-143

liquidity-constrained households especially when compared to welfare programs; this is despite dropping unemployment from the definition of liquidity-constrained households for this analysis. We find 8.7% of households who benefited from tax-cuts through the base-rate are liquidity-constrained; this reflects the fact that although liquidity-constraints will likely correlate with income (low savings households will likely have low incomes) they do not do so perfectly.

Figure 3: Liquidity-Constrained Recipients by Welfare and Tax Bands



Source: Authors Calculation through UKMOD

These findings prove the difficulty in designing fiscal stimulus to target liquidity-constrained households due the fact they only make up 30% of the population. Even tax-cuts to low-income households are imprecise in nature in achieving this aim. Instead, we can conclude from these findings that welfare programs present the most effective way to target these households as a result of their means-tested eligibility criteria closely matching what we would consider liquidity-constraints. Further, the reverse is also true, that fiscal consolidation by tax rises will not target as many liquidity-constrained households as cuts in welfare contributions will. The implications of these findings are discussed in section 6.

Covid Economics 82, 23 June 2021: 99-143

5. Findings: Macroeconomic Model

Now that we have established that fiscal policy through welfare programs can target liquidity-constrained households, we can use our macro model to illustrate why this will strengthen the effect of fiscal stimulus and consolidation. Our model does not have the capacity to consider the implications for how the stimulus is financed on our estimates. Therefore, we could assume in all instances that stimulus is money-financed by 'helicopter-drops' by the central-bank, as such financing arrangements have little to no effect on multipliers and are similar to debt-financed[10] stimulus in a zero-lower-bound environment *(Galí, 2019)*.

5.1 Scenario 1: Positive changes to welfare contributions

5.1.1 After a 3% drop in consumption (non-ZLB*)*

We begin by estimating fiscal multipliers from expansions in social security contributions during 'normal times' recessions, meaning central banks have the capacity and mandate to respond to any expansionary effects. This is simulated by a 3% fall in baseline consumption. Figure 4 shows the effects of expansionary efforts from each MPC size, with higher MPCs naturally influencing a stronger increase consumption and decrease in net-exports. This displays why the Barro-Ramsey consumption has such strong implications for the efficacy of fiscal stimulus, as an MPC of 0, as displayed, would induce no economic response.

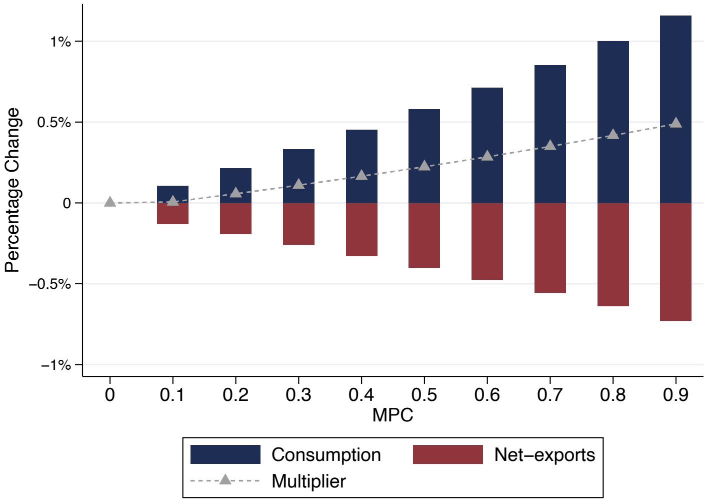
Both relationships are linear, which reflects the central bank's ability to control the expansionary effects and avoid exponential increases under a pre-determined schedule. Under this scenario, the base-rate rises from 0.4 to 0.6 as the MPC rises. Looking at just consumption, the break-even point of 1 (where governments induce more consumption than they put in) comes once all beneficiaries spend more than 80% of their stimulus. But the multiplier effect is positive, meaning that although the central bank controls the strength of the response, it does not fully offset it as the Taylor Rule construction only recommends small incremental increases according to its policy rules.

[10] See Max Corden (2010) for further detail on the long-run effects debt-financed stimulus

Covid Economics 82, 23 June 2021: 99-143

CEPR PRESS

Figure 4: Increase in Welfare Contributions on Consumption and Net-Exports by MPC Size (Scenario 1.1)



Source: Authors Calculation

Consistent with standard Keynesian theory, the effectiveness of expansionary effects is constrained through the presence of leakages from savings (1-MPC) and imports. Regarding the latter, if households spend their windfall on imported goods the gain will not be spent in the domestic economy. Although we are not aware of existing estimates of the percentage of household expenditure spent on imports, we calculate this ourselves by multiplying the household expenditure on each commodity by the import penetration of the given commodity,[11] finding 12% of household spending involves imported goods. We perform this for each household income group and find, surprisingly, no significant variation when we compare across income deciles as we would expect when looking at the expenditure of specific commodities (negative correlation between food expenditure and income). As the expansionary effects naturally result in a partial strengthening of economic conditions, we see imports become cheaper and the reverse for exports, resulting in a net loss.

[11] See Appendix B

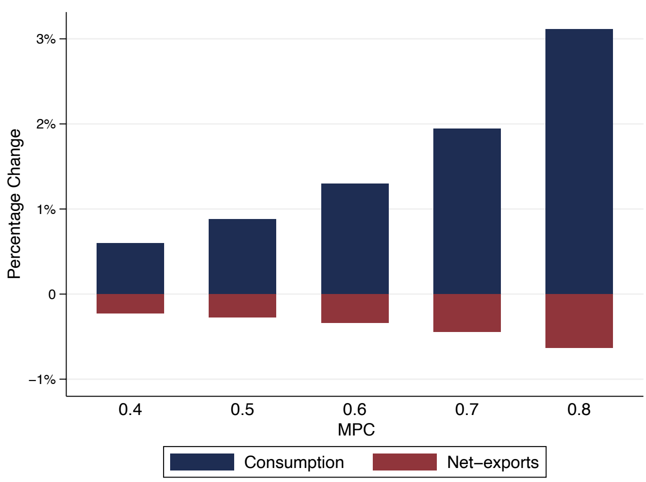
Covid Economics 82, 23 June 2021: 99-143

5.1.2 After a 5% drop in consumption (ZLB)

Scenario 1.2 looks at the above conditions but now increasing the size of recession to a 5% drop in baseline consumption to constrain interest rates at the zero-lower-bound. This is displayed in Figure 5 where we start to see a more exponential rise in output following the expansion. Many economic models suggest higher multipliers from fiscal expansions under this scenario, notably Hall (2009). We find similarly that under this scenario, not only is the consumption response greater (including a shallower fall in net exports), but the marginal increase in the multiplier is also positive as the MPC rises. This reflects the effect of idle monetary policy following an expansion from fiscal stimulus. We now focus just on the MPCs of 0.4-0.8 as these are the plausible MPC range of liquidity-constrained households. In doing so we see most of the 'heavy lifting' in terms of output increases is being done by the strong consumption response, hence why the size of the MPC is important in leveraging this output reaction.

As mentioned above, we induce a capital inflow scenario by decreasing external financial pressures. This results in a higher import response, depressing net exports and the multiplier. This finding is consistent with other papers that induce capital outflows in workhorse macro models, such as Blanchard et al (2015) who find short-run

Figure 5: Increase in Welfare Contributions on Consumption and Net-Exports by MPC Size During ZLB (Scenario 1.2)



Source: Authors Calculation

Covid Economics 82, 23 June 2021: 99-143

contractionary effects from capital inflows through a reduction in net exports from currency appreciation. The small response in net-exports reflects the central banks inability to defend the exchange rate, resulting in a smaller fall in net-exports.

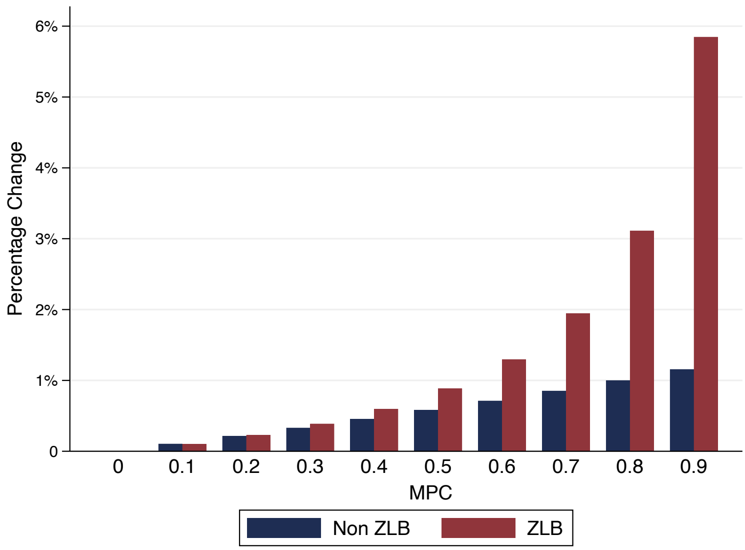
The scenarios presented are all under demand-push recessions, meaning the recession is caused by some shock to aggregate demand (in our case consumption). Of course, this is not the only form a recession can take. A 'cost-push' recession caused, for example, by a shock to the production process (such as a rise in oil prices) which increases inflation and causes a contraction in output is conceivable. This situation is often described as an impossible scenario for policy makers as measures – such as stimulus through welfare programs – can recover some of the output lost but at the cost of further increased inflation. When we simulate this potential scenario, we find no response in our output estimates – as one would expect – but although a supply shock does decrease output and increase inflation, the fiscal intervention has little further increase in inflation under this scenario than it does when there is no supply shock. Instead, we find inflation is far more sensitive to increases when inflation expectations are above the central bank target. Under both normal times and a zero-lower-bound scenario, any increase in inflation expectation translates one-for-one into inflation increases. Therefore, the starting position of inflation is important for policy makers to consider, as our results suggest the 1% intervention increases inflation by 0.8%, but inflation from stimulus often takes time to materialise due to price 'stickiness' (Galí, 2019). Therefore, it is important for the policy maker to consider existing inflation levels and a measure of future expectations. Policy makers must, as always, be cognisant of the fact that expansions can cause inflationary pressures, which during a zero-lower-bound scenario can be particularly strong.

We can show with this model the implications of transferring stimulus to this strong-demand side channel and show how the presence of zero-lower-bound interest strengthens the effect of the stimulus boost. Figure 6 compares just consumption responses under Scenario 1.1 under a normal interest setting and Scenario 1.2 which zero-lower-bound interest rates. We can see this exponential rise clearly here under the

Covid Economics 82, 23 June 2021: 99-143

latter, reflecting the highly responsive economic environment created without the presence of a monetary response.

Figure 6: Increase in Welfare Contributions on Consumption by MPC Size



Source: Authors Calculation

5.2 Scenario 2: Negative changes to welfare contributions

5.2.1 After a 5% drop in consumption (ZLB)

This second scenario assess the impacts from a fall in contributions for these liquidity-constrained households. This is in part motivated by the finding that households, especially with low savings, are more responsive to negative income shocks than positive (Bunn, et al., 2017). This decision was taken by the UK through a number of welfare reforms between 2013-16, notably through the introduction of a ceiling on the amount of welfare a household could receive, known as the 'benefit cap'.

We therefore take the above setting of a major contraction in output under a zero-lower-bound scenario, but now reverse the sign of the policy shock to test for the effects from cuts to welfare contributions. We also reverse the external financial pressure parameter to induce a small capital outflow response to the depreciation of the current account that follows, consistent with Tanner (2017).

Covid Economics 82, 23 June 2021: 99-143

Our results show, just as scenario 1.2 finds, an exponential effect during the presence of zero-lower-bound interest rates. Overall, we find the same results as in 1.2 but now with the sign reversed, resulting in a strong fall in consumption as the MPC rises and a shallow increase in net-exports due to capital outflows. Figures 5 and 6 can therefore be flipped to show the effects of cuts to contributions on different MPC sizes.

5.3 Summary

So far, we have shown the response for each MPC assumption. For the summary we make a decision about the average MPC of all those who benefit from stimulus through welfare or through tax-cuts based on the above microsimulation exercise. We found a strong presence of liquidity-constrained households within existing welfare programs, particularly means-tested, from our microsimulation exercise; therefore, it is appropriate to consider a high proportion of welfare recipients as obtaining the high MPC range specified by Canbary and Grant (2019). However, not all welfare recipients are liquidity-constrained, therefore we assume that the remaining 29% have an MPC of 0, which will lower the average MPC of all welfare recipients. We take an average from Canbary and Grant's range of 0.85 and apply it to 71% of households who benefit and assume 0 for the rest[12], resulting in average MPC of 0.59 out of positive income shocks. This can be strengthened if the policy maker directs stimulus through specific welfare programs such as housing benefit which impact's 84% of liquidity-constrained households, but for our estimates we take the average across all means-tested programs. We further apply this analysis to stimulus through tax-cuts to provide a valid comparison. We found only 8.7% of beneficiaries from cuts to the lowest 'personal allowance' tax band would be liquidity-constrained. Applying the same rules above results in an average MPC of 0.8 from tax-cuts to this tax-band; we should note that this is a lower estimate than the estimates of average MPC from the US tax cut of 0.20-0.40 (Johnson, et al., 2006).

[12] This is a strict interpretation of the Barro-Ramsey consumption model which may understate results, but we believe there is merit in providing conservative estimates

Covid Economics 82, 23 June 2021: 99-143

For cuts to contributions, we take use Bunn et al's (2017) MPC estimate for households with low net liquid assets to income ratio of around 0.9, which somewhat resembles the Zeldes definition of liquidity-constraints. Again, we apply this to 71% of means-tested welfare recipients and assume 0 otherwise, yielding an average MPC of 0.63 out of negative income shocks. Using these estimates, we simulate the effects of positive and negative changes to welfare contributions in scenario 1.2 (zero-lower-bound interest rates) and compare them to stimulus through tax-cuts, obtaining the following results presented in Table 2.

We now further solve for investment which yields some interesting results. A standard IS/LM framework suggests investment is directly proportional to savings, so we would expect as the less is saved and more consumed (as the MPC rises) investments should fall

Table 2: Multiplier Estimates from Tax-Cuts and Changes in Welfare Contributions

	Tax-Changes	**Welfare Changes**	
	Stimulus	Stimulus	Consolidation
MPC	*0.08*	*0.59*	*0.63*
Consumption	0.16	1.25	-1.46
Investment	0.04	0.59	-0.70
Net-exports	-0.16	-0.33	0.36
Multiplier	**0.05**	**1.51**	**-1.79**

as Barro (2009) argues. Our results suggest the opposite, that stimulus has a positive effect on investment and is rising with the MPC. This has been observed in reality, where the US stimulus checks improved firm level investment due to the increased profitability at the firm level following the higher economic activity (Correa-Caro, et al., 2018). Our investment equation summarises the effect of stimulus on investment into two countervailing forces: the reduction in savings increasing the real interest rate reducing investment and the improved output gap increasing it. Our New Keynesian model therefore suggests the latter force is stronger than the former, likely as a result of the zero-lower-bound environment.

Covid Economics 82, 23 June 2021: 99-143

CEPR PRESS

Overall, our show a strong multiplier effect from positive increases particularly as a result of the positive effect on consumption, yielding a positive impact multiplier of 1.51. When we compare this to tax-cuts we see a far lower estimate, with very little impact on the economy as a result of the far lower average MPC from beneficiaries. For cuts to contributions, we find a negative multiplier of 1.79 again as a result of the substantial contractionary effect on domestic consumption as a result of the higher MPC from negative income shocks. Our results are clearly sensitive to the size of the MPC, as in our model there is little difference between the positive and negative MPCs from welfare changes, but we see noticeably different outcomes. This reflects the exponential nature of Figure 6, which after an MPC of 0.5 grows rapidly. This gives further weight to the necessity of comparing this paper's results with estimates from models with greater precision.

6. Discussion

Our results have confirmed the key intuition of this paper, that welfare programs target liquidity-constrained households and subsequent changes to the contributions level of means-tested programs will has a strong effect on economic outcomes. Our model suggests every 1-unit increase in means-tested welfare will result in an increase output by 1.5 by improving both consumption and investment, whereas every 1-unit cut will result in a 1.8 fall in output. This results in two clear policy implications: the economic effects of stimulus can be maximised when directed through welfare programs, whereas fiscal consolidation from these programs will have a strong negative effect on output.

It is obvious that these results run contrary to past policy decisions by the UK government. Specifically, over the course of 2010-15 we saw a strategy of tax cuts to the top marginal rate and cuts in welfare contributions notably with the introduction of a 'benefit cap' in 2013. The latter was justified on the grounds of debt consolidation, with the stated aim to:

> *"Secure the economic well-being of the country by reducing spending on benefits." (House of Commons Work and Pensions Committee, 2019)*

Covid Economics 82, 23 June 2021: 99-143

This paper implies that the logic of this approach is inverted, as tax cuts would have not benefited many, if any, liquidity-constrained households but cuts to welfare would have, specifically by the imposition of this benefit cap. This corresponds with a popular narrative in the UK surrounding welfare, which often resorts to notions of 'economic 'cost', implying a policy trade-off between social and economic objectives.

This policy framing of welfare as a necessary social instrument during good times but an economic extravagance during recessions once the economic environment worsens, this naturally sets the program up for retrenchment (Gamble, 2016). This runs contrary to the results presented so far, that raising welfare can *contribute* to recovery, not deepen it. This paper confronts the lack of appreciation for the economic gains from welfare programs which may be leading to these counter-productive measures.

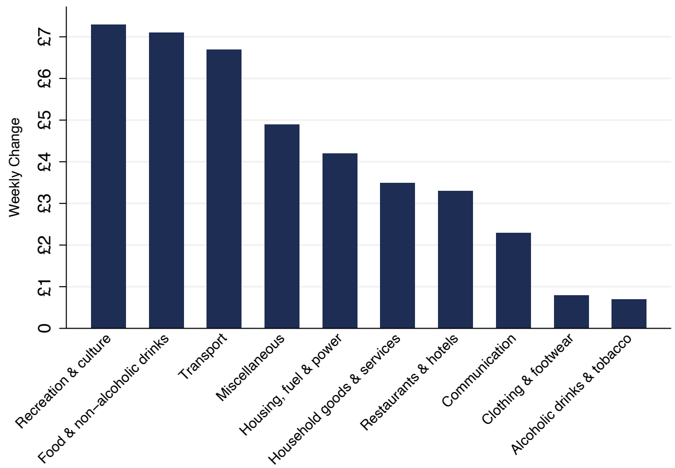
These findings have clear implications for the 2020 COVID-19 induced recession, as the Bank of England's research during the crisis showed build-up of savings in older households with higher levels of income and savings (Bank of England, May 2021), whereas households with lower savings prior to the crisis have seen rises in debt (Brewer & Handscomb, 2021). This distribution from low to high savings households with high liquidity will not result in a consumption response which is key for the UK's short-term recovery post restrictions. This paper not only echoes that concern but highlights how the £20 boost to Universal Credit could be an important source of demand in alleviating this effect. Its withdrawal in Q4 2021 could likely have strong contractionary effects on domestic output and exacerbate the macroeconomic effects of the expiry to the job-retention-scheme scheduled for the same period.

Further, we can speculate where these gains induced from Universal Credit currently support by looking at official consumption figures by household income. We can estimate the marginal gain as incomes move from the bottom decile to the second by showing the difference in weekly spend on each commodity displayed in Figure 7. This shows as incomes rise at the lower end of the income distribution, recreation & culture along with food & alcoholic are the sectors that gain the most. Coincidently, these are the sectors that have been impacted the most by COVID-19 restrictions (Brewer, et al.,

Covid Economics 82, 23 June 2021: 99-143

2021), so future demand post restrictions in these sectors could be aided if lower-income households see their incomes rise.

Figure 7: Change in Weekly Expenditure by Sector Between Bottom and Second Income Deciles



Calculated by taking the difference in weekly expenditure by COICOP sector between the bottom and second income decile. Source: ONS (2019)

As this paper intends to serve as a benchmark for future research, there are number of areas we recommended further explorations into. First and foremost, we have written at length regarding what we can and cannot infer from our macroeconomic model. As such a first exercise would be to cross-check these results against others more sophisticated DSGE/(s)VAR methods, in a similar vain to Pappa et al (2015). These methods may provide a more accurate account of this position and will have the further advantage of comparing these results over time. Testing this papers intuition against these robust frameworks would be an important exercise in upgrading our economic 'illustration' to providing precise forecasts. An improvement could also be made into our microsimulation efforts to identify liquidity-constrained households. Specifically, with the introduction of survey data in credit-constraints, similar to Dolls et al (2012), with the introduction of data from the FCA dataset mentioned earlier. Our dataset is only from 2018, so this could further be applied to multiple years to assess how our results change over time, specifically if the onset of the 2020 recession changed the number of

Covid Economics 82, 23 June 2021: 99-143

CEPR PRESS

liquidity-constrained households in the population and their presence in welfare programs.

Similarly, studies that can observe this in reality could provide confirmation on the causal linkage between welfare contributions and economic outcomes. Specifically, the roll-out of Universal Credit – which has acted as a virtual cut to contributions (IFS, 2019) – was initially rolled-out according to randomly chosen 'NUTS 3' regions between 2015-2018 (DWP, 2018). This presents the opportunity for a natural experiment into the effects this has had on regional economic outcomes, notably consumption. However, to date there is no consumption data that corresponds to 'NUTS 3' regions, only the larger 'NUTS 2' level, making causal inference near impossible. If this data, or if other indicators of economic activity becomes available, there may be an opportunity to test for these effects in this paper regarding the contractionary effects of welfare cuts on regional economic outcomes.

Lastly, as the assumption of this paper is that fiscal stimulus cannot target liquidity-constrained households in part due to administrative challenges, there may be an opportunity for Central Bank Digital Currencies (CBDCs) to meet this challenge. At the time of writing CBDCs are still being developed by the Bank of England and the Treasury, but these may offer an opportunity for future monetary and/or fiscal stimulus if it has the capability to collect data on financial assets. This is a programmability choice that it designs can take, which would allow for governments and central banks to put in place the infrastructure necessary to be able to identify and transfer cash to liquidity-constrained households (Dyson & Hodgson, 2016). Therefore, the logic of this paper can be extended to an exploration into how CBDCs can assist stimulus designs and minimise traditional trade-offs between efficacy and feasibility that this paper has sought to answer.

7. Conclusion

Our paper began by explaining that in the presence of zero-lower-bound interest rates, fiscal policy can and must take a frontline role as an economic stability mechanism, but its effectiveness is constrained so long as it currently has no way to target liquidity-

Covid Economics 82, 23 June 2021: 99-143

constrained households specifically. The goal of this paper was to therefore determine if existing welfare programs can meet this challenge by identifying the proportion of these households already claiming welfare and illustrating the implications of changes to contributions levels with a macroeconomic model. We found a strong presence of liquidity-constrained households within these programs, specifically those that require means-testing. This implies two conclusions; first governments can use these means-tested welfare programs as an efficient means of transferring cash-based fiscal stimulus those households who are liable to spend it. Our macroeconomic model estimated a strong impact fiscal multiplier of 1.5 from increases in these contributions. The second conclusion is that cuts to these programs will have a stronger than usual contractionary response on domestic consumption if the households that will see their income fall are even more sensitive to such reductions in their income than they are to positive changes. Our model estimated negative multipliers of 1.8, suggesting that the cuts to welfare contributions, that formed a key part of the UK's austerity measures, would have likely had a strong contractionary effect on domestic output. This paper therefore offers a unique contribution to fiscal policy literature by identifying the role existing welfare programs play in strengthening their effects on macroeconomic outcomes.

Bibliography

Agarwal , S. & Qian, W., 2014. Consumption and Debt Response to Unanticipated Income Shocks: Evidence from a Natural Experiment in Singapore. *American Economic Review,* 104(12), pp. 4205-30.

Bank of England, 2016. *Inflation Report November*

Bank of England, May 2021. *Monetary Policy Report ,* London

Covid Economics 82, 23 June 2021: 99-143

Barro, R. J., 1974. Are Government Bonds Net Wealth. *Journal of Political Economy,* 82(6), pp. 1095-1117.

Barro, R. J., 1981. Output Effects of Government Purchases. *The Journal of Political Economy,* 89(6), pp. 1086-1121.

Barro, R. J., 2009. *Demand Side Voodoo Economics,* Berkeley: The Berkeley Electronic Press.

Barro, R. J., 2010. *The Stimulus Evidence One Year On.* Wall Street Journal.

Blanchard , O., 2009. The State of Macro. *Annual Review of Economics,* Volume 1, pp. 209-228.

Blanchard, O., Ostry, J. D., Ghosh, A. R. & Chamon, M., 2015. Are Capital Inflows Expansionary or Contractionary? Theory, Policy Implications, and Some Evidence. *International Monetary Fund, Washington*, p. IMF Working Paper WP/15/226.

Blanchard, O. & Perotti, R., 2002. An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *Quarterly Journal of Economics Volume,* 117(4), pp. 1329-1368.

Bourquin, P. & Waters, T., 2019. *The effect of taxes and benefits on UK inequality,* London: IFS Briefing Note BN249.

Brewer, M., Corlett, A., Handscomb, K. & Tomlinson, D., 2021. *The Living Standards Outlook ,* London: Resolution Foundation.

Brewer, M. & Handscomb, K., 2021. *The debts that divide us,* London: Resolution Foundation.

Bunn, P., Le Roux, J., Reinold, K. & Surico, P., 2017. The consumption response to positive and negative income changes. *Bank of England: Staff Working Paper No. 645*.

Campbell, J. R. & Hercowitz, Z., 2019. Liquidity Constraints of the Middle Class. *American Economic Journal: Economic Policy,* 11(3), pp. 130-55.

Canbary, Z. & Grant, C., 2019. The Marginal Propensity to Consume for Different Socio-economic Groups. *Economics and Finance Working Paper Series: Working Paper No. 1916*, October.

Christiano, L., Eichenbaum, M. & Rebelo, S., 2010. *When is the Government Spending Multiplier Large?,* Evanston: Northwestern University.

Cochrane, J., 2010. *Fiscal Stimulus, RIP*

Corden , M. W., 2010. The theory of the fiscal stimulus: how will a debt-financed stimulus affect the future?. *Oxford Review of Economic Policy,* 26(1), pp. 38-47.

Covid Economics 82, 23 June 2021: 99-143

Correa-Caro, C., Medina, L., Poplawski-Ribeiro, M. & Sutton, B., 2018. Fiscal Stimulus Impact on Firms' Profitability During the Global Financial Crisis. *IMF Working Paper WP/18/251*.

Crossley, T. F., Fisher, P., Levell, P. & Low, H., 2021. *MPCs through COVID: spending, saving and private transfers,* London: Institute for Fiscal Studies.

Dolls, M., Fuest, C. & Peichl, A., 2012. Automatic stabilizers and economic crisis: US vs. Europe. *Journal of Public Economics,* Volume 96, p. 279–294.

Dolls, M., Fuest, C. & Peichl, A., 2012. Automatic stabilizers and economic crisis: US vs. Europe. *Journal of Public Economics,* 96(3-4), pp. 279-294 .

DWP, 2018. *Universal Credit Transition Rollout Schedule ,* Department for Work and Pensions.

DWP, 2021. *How to claim Universal Credit: step by step.* [Online]
Available at: https://www.gov.uk/universal-credit/eligibility

Dyson, B. & Hodgson, G., 2016. *Digitial Cash: Why Central Banks Should Start Issuing Electronic Money,* Positive Money.

Edelberg, W., Eichenbaum, M. & Fisher, J., 1999. Understanding the Effects of a Shock to Government Purchases. *Review of Economic Dynamics,* 2(1), pp. 166-206 .

Erceg, C. J. & Lindé, J., 2014. Is There a Fiscal Free Lunch in a Liquidity Trap?. *Journal of the European Economic Association,* 12(1), p. 73–107.

FCA, 11 February 2021. *Financial Lives 2020 survey: the impact of coronavirus,* London

Fisher, J., Johnson, D. & Smeeding, T., 2019. Estimating the Marginal Propensity to Consume Using the Distributions of Income, Consumption, and Wealth. *Federal Reserve Bank of Boston Research Department Working Papers No. 19-4*.

Galí, J., 2019. The effects of a money-financed fiscal stimulus. *Economics Working Paper Series: Working Paper No. 1441*, July.

Gamble, A., 2016. The Battle of Ideas. In: *Can the Welfare State Survive?.* Cambridge: Polity Press, pp. 22-29.

Gechert, S., Paetz, C. & Villanueva, P., 2021. The macroeconomic effects of social security contributions and benefits. *Journal of Monetary Economics,* Volume 117, pp. 71-584.

Goldby, M., Laureys, L. & Reinold, K., 2015. *An estimate of the UK's natural rate of interest,* Bank Underground.

Gross, T., Notowidigdo, M. J. & Wang, J., 2016. The Marginal Propensity to Consume Over the Business Cycle. *NBER Working Paper No. 22518*, August.

Hall, R. E., 2009. By How Much Does GDP Rise if the Government Buys More Output?. *Brookings Papers on Economic Activity,* Issue 2, pp. 183-249.

Hall, R. E. & Mishkin, F. S., 1982. The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households. *Econometrica,* 50(2), pp. 461-481.

HM Revenue & Customs, 2020. *Increase to Working Tax Credits - what this means.* [Online] Available at: https://www.gov.uk/government/news/increase-to-working-tax-credits-what-this-means

House of Commons Work and Pensions Committee, 2019. *The benefit cap: Twenty-Fourth Report of Session 2017–19,* London: House of Commons .

IFS, 2019. *Universal credit and its impact on household incomes: the long and the short of it,* London: Institute for Fiscal Studies.

Ilzetzki, E., Mendoza, E. G. & Végh, C. A., 2013. How big (small?) are fiscal multipliers?. *Journal of Monetary Economics,* 60(2), pp. 239-254 .

Ipsos MORI, January 2021. *The Health Foundation COVID-19 Survey – third poll,* The Health Foundation.

Jappelli, T., Pischke, J.-S. & Souleles, N., 1998. Testing For Liquidity Constraints In Euler Equations With Complementary Data Sources. *The Review of Economics and Statistics,* 80(2), pp. 251-262.

Jappelli, T. & Pistaferri, L., 2014. Fiscal Policy and MPC Heterogeneity. *American Economic Review,* 6(4), pp. 107-36.

Jappelli, T. & Pistaferri, L., 2014. Fiscal Policy and MPC Heterogeneity. *American Economic Journal: Macroeconomics,* 6(4), pp. 107-36.

Johnson, D. S., Parker, J. A. & Souleles , N. S., 2001. Household Expenditure and the Income Tax Rebates of 2001. *American Economic Review,* 96(5), pp. 1589-1610.

Johnson, D. S., Parker, J. A. & Souleles, N. S., 2006. Household Expenditure and the Income Tax Rebates of 2001. *American Economic Review,* 96(5), pp. 1589-1610.

Koenig, G. & Myles, A., 2013. *Social Security's Impact on the National Economy,* Washington: AARP Public Policy Institute.

Krugman, P., 2000. How Complicated Does the Model Have To Be?. *Oxford Review of Economic Policy,* 16(4), pp. 33-42.

Nakamura, E. & Steinsson, J., 2011. Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions. *Working Paper Series (17391)*.

OBR, 2020. *Potential output and the output gap.* [Online] Available at: https://obr.uk/forecasts-in-depth/the-economy-forecast/potential-output-and-the-output-gap/#outputgap

Covid Economics 82, 23 June 2021: 99-143

Office for Natonal Statistic, 2019. *Household expenditure as a percentage of total expenditure by disposable income decile group: Table A8.* [Online]
Available at:
https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/expenditure/datasets/householdexpenditureasapercentageoftotalexpenditurebydisposableincomedecilegroupuktablea8

Parker, J. A., Souleles, N. S., Johnson, D. S. & McClelland, R., 2013. Consumer Spending and the Economic Stimulus Payments of 2008. *American Economic Review,* 103(6), pp. 2530-53.

Payne, S. & Parker, G., 2021. *Boris Johnson faces anger from Tory MPs over planned welfare benefit cut,* s.l.: Financial Times.

Ramey, V., 2009. Identifying Government Spending Shocks: It's All in the Timing. *NBER Working Paper No. 15464*, October.

Ramey, V. A. & Shapiro, M. D., 1998. Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference Series on Public Policy,* Volume 48, pp. 145-194.

Romer, C. D. & Romer, D. H., 2010. The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. *American Economic Review,* Volume 100, p. 763–801.

Romer, P., 2016. *The Trouble With Macroeconomics,* s.l.: Stern School of Business, New York University.

Runkle, D. E., 1991. Liquidity constraints and the permanent-income hypothesis: Evidence from panel data. *Journal of Monetary Economics,* 27(1), pp. 73-98.

Sargent, T. J., 1987. Macroeconomic theory. In: 2nd ed. Boston: Academic Press, p. 397–398.

Shoag, D., 2013. Using State Pension Shocks to Estimate Fiscal Multipliers since the Great Recession. *American Economic Review,* 103(3), pp. 121-124.

Spinnewijn, J., 2020. The Trade-off between Insurance and Incentives in Differentiated Unemployment Policies. *Fiscal Studies,* 41(1), pp. 101-127.

Tamegawa, K., 2012. Closed Form of Fiscal Multipliers in a DSGE model with Liquidity-Constrained households. *Economics Bulletin,* 32(4), pp. 3148-3157.

Tsurugaa, T. & Wake, S., 2019. Money-financed fiscal stimulus: The effects of implementation lag. *Journal of Economic Dynamics & Control,* Volume 104, pp. 132-151.

Zeldes, S. P., 1989. Consumption and Liquidity Constraints: An Empirical Investigation. *Journal of Political Economy,* 97(2).

Appendix A

Consumption

Our model starts from the standard 'textbook' consumption function, we differentiate autonomous (α_{C0}) and induced (α_{CY}) consumption dynamics; alternatively. Taxes are defined as a constant tax share τ and a one-off lump sum tax component TP_t.

$$C_t = \alpha_{C0} + \alpha_{CY}(Y_t(1-\tau) - TP_t$$

Such a simple equation, for instance, cannot capture the differential effects of households who are able to follow the permanent income hypothesis (PIH) compared to liquidity-constrained households; this is the key consideration of this paper. Therefore, we extend the following functions to include these effects. Autonomous consumption is now defined as:

$$\alpha_{C0} = Y^P * [(1-\sigma) - (1-\sigma_{CY})](1- \tau)$$

Here we include long and short run savings rates of σ and σ_{cyc} respectively. This allows α_{CY} to be interpreted as the pre-tax short run marginal propensity to consume $(1-\sigma_{CY})$ out of potential output Y^P. Therefore, we can reinterpret α_{C0} as:

$$\alpha_{C0} = Y^P * (1-\sigma)(1- \tau) = Y^P * (1-\tilde{\sigma})$$

Where $(1-\tilde{\sigma}) = (1-\tilde{\sigma}) * (1-\tau)$, therefore taking the form of long run savings adjusted for tax. Therefore, the linear equation can now be presented as:

$$C_t = (1- \tilde{\sigma})Y^P + \left(1-\sigma_{cyc}\right)\{[(1- \tau)(Y_t - Y^P)] - TP_t\}$$

Where $(1- \tilde{\sigma})Y^P$ is our long run APC component out of potential output and $\{[(1- \tau)(Y_t - Y^P)] - TP_t\}$ is our short run component. Strict interpretation of the PIH and/or Ricardian Equivalence (REH) suggests $(1-\sigma_{CY})$=0 as argued by John Cochrane (Cochrane, 2010). Under this argument negative TP_t (any tax cut/transfer form of fiscal policy) will have no effect. This fundamental assumption of this paper is

Covid Economics 82, 23 June 2021: 99-143

that $(1-\sigma_{CY})$>0 for welfare recipients, therefore negative TP_t will induce a consumption response.

Investment

We start as above with an extension to a simple framework. We start by assuming a natural real rate of interest $(\bar{r})$ that would hold a zero-output-gap at a steady state of marginal capital to depreciation.

$$I_t = \tilde{\alpha}_{I0} + \alpha_{Ir}(r_t - \bar{r})$$

Where $\tilde{\alpha}_{I0}$ represents autonomous investment in terms of potential output. For now, we assume investment exactly equals savings so that $\tilde{\alpha}_{I0}$ must equal: $\psi Y^P = \sigma * (1-\tau)Y^P$. Where ψ is the depreciation of the capital stock at the steady state of investment.

Net Exports

In extending the above to include external shocks we begin by including international trade. The level of both exports and imports is assumed to comprise both a long run component (as a constant fraction of potential output) and short run component which is dependent on deviations of relative price of exports from long-run trends:

$$X_t = Y^P[x + \eta_x * rpx_t]$$

$$\mathrm{IM} = Y^P[im + im_{cyc} * gap_t + \eta_{im} * rpim_t]$$

Here, x and IM denote exports and imports respectively, these are determined by long-run external prices and productivity levels in the relevant traded good sectors. For imports we include both long run import share im (determined by long-run external prices and productive capacity) and im_{cyc} as a short-run marginal propensity to import[13] along with the output gap $gap_t = \left(\frac{Y_t}{Y^P} - 1\right)$. The η parameters are a response function (assumed to be >0) to rpx_t and $rpim_t$ which are the percentage deviation of

[13] See appendix for how we calibrate this parameter

Covid Economics 82, 23 June 2021: 99-143

the relative price of exports and imports to their long-run parameters RPX_t $RPIM_t$defined as:

$$RPX_t = \left[\frac{S_t * P_t^X}{P_t}\right]$$

$$RPIM_t = \left[\frac{S_t * P_t^{IM}}{P_t}\right]$$

S_t is the nominal exchange rate, P_t^X and P_t^{IM} are the world currency price of exports and imports, while P_t is the domestic price level. The short run rpx_t and $rpim_t$ parameters are therefore determined by the real exchange rate q_t and the scaled external terms of trade TT_t which is simply $TT_t = P_t^X / P_t^{IM}$:

$$rpx_t = q_t + (1 - \nu)\ln(TT_t)$$

$$rpim_t = q_t - \nu \ln(TT_t)$$

Where ν is the relative importance of imports, q_t is determined by:

$$Q_t = \left[\frac{S_t * P_t^{EXT}}{P_t}\right]$$

P_t^{EXT} corresponds to level of external prices which can be written as a weighted average of export and import prices:

$$P_t^{EXT} = (P_t^x)^\nu (P_t^{IM})^{(1-\nu)}$$

We then extend this model to incorporate domestic and foreign components. First, we create a baseline using $r^{*\,EXT}$ as the sum of the external natural rate of interest and RP as the risk premium, such that $r^* \ = r^{*\,EXT} + rp^*$ so that the steady state natural interest rate converges towards to the steady state marginal product of capital net of

Covid Economics 82, 23 June 2021: 99-143

depreciation. This allows us to summarise external financial pressures as the divergence between external interest rates plus risk premium to their baseline values, such that $efp_t = [r_t^{EXT} + rp_t] - [r^{*\,EXT} + RP^*]$. In assuming that domestic and external financial pressures are symmetric (e.g., tight domestic monetary policy will lead to a fall in the relative prices of exports and imports through real exchange rate appreciation) we can rewrite the above equation so that:

$$rpx_t = q_t + (1-\nu)\ln(TT_t) = [r^* - r_t] + efp_t + (1-\nu)\ln(TT_t)$$

$$rpim_t = q_t - \nu\ln(TT_t) = [r^* - r_t] + efp_t - \nu\ln(TT_t)$$

Where $[R^* - R_t]$ represents domestic monetary policy, $efp_t + (1-\nu)\ln(TT_t)$ represents external shocks. We can therefore rewrite the first equation so that export (supply) and import (demand) equations can be written as[14]

$$X_t = Y^P[x + \eta_x * (r^* - r_t) + \eta_x efp_t + \eta_x(1-\nu)\ln(TT_t)]$$

$$\mathrm{IM} = Y^P[im + im_{cyc} * gap_t + \eta_{im} * (r^* - r_t) + \eta_x efp_t - \eta_x \nu\ln(TT_t)]$$

<u>Monetary Response</u>

Assuming a Taylor Rule construction, monetary policy takes the following form:

$$r^{NOM} = r^{NAT} + \pi^e + \beta_\pi(\pi - \pi^t) + \beta_{gap}(gap) + \beta_{efp}(efp_t) + r^{disc}$$

A clear assumption here is that we assume π^e exogenous which allows us to compute output gaps without more computationally intensive frameworks. We include efp_t to incorporate the typical considerations of international financial shocks on the central Bank's setting of interest rates in an open economy. Inflation is determined by according to traditional Phillips Curve framework:

[14] Green denotes exogenous variables and red for exogenous

$$\pi = \pi^e - \frac{1}{\eta_{SRAS}}(gap - SS) + \theta efp_t$$

We determine short-run elasticity of supply by η_{SRAS}, θ represents the fraction of efp_t that impacts the real exchange rate. As is important for this paper, the inverse of the above setting represents the real interest rate when nominal interest rates are at the zero-lower-bound (ZLB). The equilibrium of real interest rates is therefore a combination of above:

$$r_t = r^{NAT} + \beta_\pi(\pi^e - \pi^t) + \beta_{gap}(gap) + \beta_{ss}(ss_t) + \beta_{efp}(efp_t) + r^{disc}$$

The response parameters determine the strength of the Bank's response to a given component, where:

$$\beta_{gap} = \frac{\beta_\pi - 1}{\eta} + \beta_{gap}$$

CEPR PRESS

Appendix B

I calibrate im_{cyc} by multiplying the average weekly spend of welfare recipients on each commodity (ONS Data) by the import penetration of the given commodity. Formally:

$$im_{cyc} = \sum_{d=1}^{D} \left(Yd \sum_{i=1}^{i} C_{id} \cdot IM_{id} \right)$$

Here, we sum weekly spend Yd by the given commodity C_{id} and multiply that given commodity by the import penetration IM_{id}. We do this for each income decile d to check for any variation across income groups. There may be imprecisions in this approach, as we assume the percentage spend on each commodity involves an equal proportion of imported goods.

Appendix C

Figure 4: Scenario 1.1

MPC	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Consumption	0.00	0.11	0.22	0.33	0.45	0.58	0.71	0.85	1.00	1.16
Investment	0.00	0.03	0.04	0.04	0.04	0.04	0.05	0.05	0.06	0.06
Net-exports	0.00	-0.13	-0.19	-0.26	-0.33	-0.40	-0.48	-0.56	-0.64	-0.73
Multiplier	0.00	0.01	0.06	0.11	0.17	0.22	0.28	0.35	0.42	0.49

Figure 5: Scenario 1.2

MPC	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Consumption	0.00	0.10	0.23	0.39	0.60	0.88	1.30	1.94	3.11	5.85
Investment	0.00	0.01	0.08	0.16	0.26	0.41	0.62	0.95	1.54	2.93
Net-exports	0.00	-0.15	-0.17	-0.19	-0.23	-0.27	-0.34	-0.44	-0.63	-1.07
Multiplier	0.00	-0.03	0.14	0.35	0.63	1.02	1.58	2.45	4.03	7.71

Figure 6: Consumption Changes

MPC	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Non ZLB	0.00	0.11	0.22	0.33	0.45	0.58	0.71	0.85	1.00	1.16
ZLB	0.00	0.10	0.23	0.39	0.60	0.88	1.30	1.94	3.11	5.85

Covid Economics 82, 23 June 2021: 99-143