DISCUSSION PAPER SERIES

DP17700

FIFTY SHADES OF QE REVISITED

Martin Weale and Tomasz Wieladek

INTERNATIONAL MACROECONOMICS AND FINANCE AND MONETARY ECONOMICS AND FLUCTUATIONS

CEPR

ISSN 0265-8003

FIFTY SHADES OF QE REVISITED

Martin Weale and Tomasz Wieladek

Discussion Paper DP17700
Published 27 November 2022
Submitted 18 November 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- International Macroeconomics and Finance
- Monetary Economics and Fluctuations

FIFTY SHADES OF QE REVISITED

Abstract

Fabo, Jancokova, Kempf and Pastor (2021) use OLS regressions to show that central bankers report quantitatively larger effects of QE on output and inflation than academic researchers. They also show that central bankers are more likely to report economically/statistically significant results, advance faster in their careers and use more positive sentiment to describe their results. We reject the null hypothesis of a Gaussian distribution of the residuals in many of these specifications. We then repeat the analysis with regression estimators that are robust to a non-Gaussian residual distribution where this is feasible. We use the median (50% quantile) regression estimator and the MS regression estimator of Maronna and Yohai (2000). With these robust regression approaches, the null hypothesis that central bank and academic researchers report the same quantitative effect of QE on output and inflation cannot be rejected, with point estimates which are less than half as large. There is no evidence that the remaining results are affected by non-Gaussianity. In particular, the sentiment regressions, for which there is no evidence of non-Gaussianity, are robust to all of the estimators explored here.

JEL Classification: E52, E58

Keywords: N/A

Martin Weale - martin.weale@kcl.ac.uk
King's Business School

Tomasz Wieladek - tomaszwieladek@gmail.com
T. Rowe Price, King's College London, CEPR and CEPR

Acknowledgements
We are very grateful to Brian Fabo, Martina Jancokova, Elisabeth Kempf and Lubos Pastor for sharing their data, support and very helpful feedback.

Fifty Shades of QE Revisited

Martin Weale[(1)] and Tomasz Wieladek[(2)]
November 2022

ABSTRACT

Fabo, Jancokova, Kempf and Pastor (2021) use OLS regressions to show that central bankers report quantitatively larger effects of QE on output and inflation than academic researchers. They also show that central bankers are more likely to report economically/statistically significant results, advance faster in their careers and use more positive sentiment to describe their results. We reject the null hypothesis of a Gaussian distribution of the residuals in many of these specifications. We then repeat the analysis with regression estimators that are robust to a non-Gaussian residual distribution where this is feasible. We use the median (50% quantile) regression estimator and the MS regression estimator of Maronna and Yohai (2000). With these robust regression approaches, the null hypothesis that central bank and academic researchers report the same quantitative effect of QE on output and inflation cannot be rejected, with point estimates which are less than half as large. There is no evidence that the remaining results are affected by non-Gaussianity. In particular, the sentiment regressions, for which there is no evidence of non-Gaussianity, are robust to all of the estimators explored here.

Keywords: Monetary Policy, Quantitative Easing, Robust regression.
***JEL* classification**: E52, E58, E31.

(1) King's College, University of London. Email: martin.weale@kcl.ac.uk

(2) King's College, University of London, T Rowe Price and CEPR. Email: tomaszwieladek@gmail.com

Acknowledgements
We are very grateful to Brian Fabo, Martina Jancokova, Elisabeth Kempf and Lubos Pastor for sharing their data, support and very helpful feedback.

1 Introduction

Fabo, Jancokova, Kempf and Pastor (2021) provide the first comparison of the effects of QE reported by central bank[1] and academic researchers. They find that central bank researchers report larger inflation and output effects of QE than researchers in academia. Central bank researchers are more likely to report significant results, derive a career benefit from their work and use more positive language to describe their findings. Fabo et al (2021) provide a great service to the profession by starting this debate, collecting these data and making them publicly available. Although the authors show that their results are robust to many perturbations, the underlying econometric tool used in their study is OLS regression. They use robust standard errors, but these do not address the distortions that can arise when this approach is applied to data with outliers.

We show that the residuals of most of their regressions have values of skewness and kurtosis which are inconsistent with a standard Gaussian distribution, except for the case of sentiment (positive language) regressions. Applying the OLS estimator in these circumstances can lead to biased estimates. We revisit their analysis with regression estimators which are robust to residuals with a non-Gaussian distribution. Once these estimators are adopted, the null hypothesis that central bank and academic researchers report the same inflation and output effects of QE cannot be rejected in most specifications. There is no evidence that the results on significance reporting or career progression are affected by non-Gaussianity. Their findings on sentiment are shown to be robust. The following section of this comment describes the methodology, while section three presents our results. Section four concludes.

[1] For full disclosure, our paper (Weale and Wieladek, 2016) was published while we worked at the Bank of England and is included in the study by Fabo *et al*. (2021). Haldane et al (2016) and Wieladek et al (2016) are included papers written by Tomasz Wieladek.

2 Data and Methodology

Fabo *et al.* (2021) collect the output and inflation effects of QE across countries from 54 different studies of QE. They also collect information on the authors' affiliations, their experience and career outcomes. Their data are clearly presented in the appendix of their paper, providing a significantly higher standard of transparency than found with most empirical work in economics. We use the data from appendix A in their paper. In table 1, we replicate the summary statistics table (table 1) of their paper. All the statistics match those reported in their paper[2]. However, we also add the fourth moment, kurtosis, to their table. For the normal distribution, kurtosis should normally take the value of 3. The estimates of kurtosis shown in table are clearly greater than 3 in the majority of cases, which means that these variables are lepto-kurtic- that is they have much fatter tails, with higher probability of outliers, than would be expected from a normal distribution. Of course, what matters is whether the lepto-kurtic nature of the dependent variables translates into lepto-kurtic (non-Gaussian) residuals. This is what we investigate below.

Fabo *et al.* (2021) rely on OLS regression to test their main hypotheses of interest. Since sometimes several QE estimates come from the same paper, the authors cluster the standard errors by paper. Furthermore, the standard errors are calculated from bootstrapped residuals with 10,000 replications of the wild bootstrap. The authors are meticulous in examining the robustness of their results to many plausible perturbations. However, their analysis relies on the OLS estimator throughout. Below we show that the distribution of residuals from their regressions have very high levels of kurtosis, even when the one or two most influential observations are removed. In most of those instances where kurtosis is closer

[2] With one very slight difference that we attribute to rounding.

to the expected value of 3, the degree of skewness implies values of the Jarque-Bera statistic which point to rejection of the null hypothesis of a Gaussian distribution in the residuals. This characteristic of the residuals in Fabo *et al.* (2021) suggests that the OLS results could be excessively influenced by outliers. We re-estimate the regressions in Fabo *et al.* (2021) with estimators which are robust to outliers and fat tails.

Table 1
Effects of QE on output and inflation by Central Bank Affiliation – Summary Statistics

	All	CB	Not CB
Panel A: Effect on Output			
Peak effect on output	1.57	1.75	1.00
	(1.25)	(1.54)	(1.00)
	[4.26]	[4.21]	[2.29]
Standardised peak effect on output	0.24	0.28	0.11
	(0.16)	(0.18)	(0.10)
	[11.02]	[7.87]	[2.29]
Cumulative effect on output	0.87	1.06	0.48
	(0.40)	(0.42)	(0.05)
	[6.34]	[5.53]	[2.92]
Standardised cumulative effect on output	0.14	0.18	0.04
	(0.04)	(0.06)	(0.01)
	[17.36]	[11.58]	[3.05]
Panel B: Effect on Output			
Peak effect on inflation	1.42	1.79	0.54
	(0.93)	(1.17)	(0.40)
	[7.54]	[5.78]	[2.19]
Standardised peak effect on inflation	0.19	0.24	0.05
	(0.11)	(0.15)	(0.04)
	[18.59]	[13.95]	[2.53]
Cumulative effect on inflation	0.89	1.35	-0.21
	(0.75)	(0.82)	(0.14)
	[14.02]	[8.06]	[9.82]
Standardised cumulative effect on inflation	0.12	0.18	-0.01
	(0.08)	(0.11)	(0.01)
	[19.15]	[17.40]	[9.30]

Note: This table reports the mean, medians (in parenthesis) and kurtosis [in brackets] for the estimated effects of QE on output and inflation, separately for papers with and without CB-affiliated authors. Papers with no "full" central banker are excluded from the statistics for authors with central banking affiliations.

There is an important difference in the bootstrapping approach in this paper and that of Fabo et al (2021). They use the wild bootstrap (Wu, 1986). The advantage of this method is that the bootstrapped sample will resemble the actual bootstrap, even if cluster sizes vary a lot. However, the wild bootstrap can be applied only to OLS and related models. We rely on the pairs cluster bootstrap (Bertrand, Duflo and Mullainathan, 2004) in this paper. While this approach doesn't work as well if cluster sizes vary a lot, it can be applied to any regression model, not just OLS. This allows us to explore whether their results are robust to regression models which are more resilient to outliers than OLS. A key question is whether the difference in bootstrap makes a difference to the results we present here. We show that, when applied to the OLS estimates, the pairs cluster bootstrap generates very similar results to the wild bootstrap, which we present in appendix A. This implies that the results in this paper are not due to differences in bootstrap approach with Fabo et al (2021).

There are three types of outliers which can affect both the estimate and inference in the standard OLS regression framework (see Rousseeuw and Leroy (2003) for more detail). Good leverage points are outliers which are on the regression line, but far away relative to all the other observations. Good leverage points affect only inference and not OLS estimates. On the other hand, observations, which are away from the regression line in y space only, referred to as vertical outliers, affect OLS estimates. Similarly, observations which are outliers in x space, referred to as bad leverage points, affect OLS estimates as well.

Edgeworth (1887) provided a means of addressing outliers which affect OLS estimates, by introducing the least absolute deviation regression estimator. Rather than minimising the sum of squares of the residuals, this estimator minimises the sum of absolute deviations of the residuals. The easiest implementation of this approach is as a median (50% quantile) regression. As the OLS estimator minimises the sum of the squared residuals, any outlying observation will get a large weight, and the greater the outlier, the greater the weight put on this observation. In contrast, the median regression estimator minimises the sum of the absolute errors, putting an equal weight on each observation. However, the caveat of this estimator is that it protects only against vertical outliers, but not against bad leverage points. This estimator also has lower efficiency than the OLS estimator if the distribution of errors is Gaussian, but generally higher efficiency in the case of a non-Gaussian distribution (Koenker and Basset, 1978).

An alternative class of robust estimators builds on the idea of using loss functions other than the OLS square and absolute deviation (median regression) estimator. This idea was initially advocated by Huber (1964) who proposed the M estimator. The latest evolution of the approach is the MM estimator proposed by Yohai (1987). The advantage of this estimator is that it has both high efficiency if the underlying distribution is Gaussian and a high breakdown point of 50%, meaning that this estimator resists contamination if outliers comprise up to 50% of observations in the sample. Furthermore, this estimator is robust to all three types of outliers: vertical outliers, good leverage points and bad leverage points. However, when dummy variables are present in the specification, Verardi and Croux (2008) recommend the MS-estimator of Maronna and Yohai (2000). We use the MM-estimator for the specifications without dummy variables and rely on the MS-estimator where dummy variables are present in the regression specification.

In the regressions presented by Fabo et al (2021), standard errors are clustered by paper at regression level and also bootstrap clustered by paper. This treatment of standard errors is possible only with the OLS and median estimators. Only bootstrap clustering is possible with the MM/MS-estimators. In Appendix A2 we show that results from OLS and quantile regressions with bootstrap clustering only are very similar to those with double clustering. This suggests that the lack of regression level clustering is unlikely to make a significant difference to the results reported here.

3 Results

Our first task is to replicate the results of Fabo *et al.* (2021) with the OLS estimator. As in their paper, the residuals are clustered by paper and bootstrapped with 10,000 replications, but we rely on the pairs cluster, rather than the wild, bootstrap. The results are shown in tables 2 and 3, for output and inflation effects as dependent variables, respectively.

For ease of comparison, we use the same labelling and presentation as in the original paper by Fabo *et al.* (2021). Tables 2 and 3 show results with the pairs cluster bootstrap, while tables A_2 and A_3 in appendix A1 show results with the wild cluster bootstrap. The p-values in tables 2 and A_2 (the output regressions) are very similar, apart from specification 3 for the cumulative effect in panel B. For the standardised effects in tables 3 and A_3 (the inflation regressions), the p-values show statistical significance at less than 1% with the wild bootstrap (table A_3), but only less than 3% with the pairs cluster bootstrap (table 3). While this shows that there is some systematic difference between these two different bootstrap approaches, this difference is small and inconsequential for inference: A statistically significant effect is detected across the board in either case. Mackinnon and Webb (2017) argue that the pairs bootstrap deviates significantly from the wild bootstrap, only when the variation across clusters is large. Since this isn't the case in these data, and the OLS results with the pairs bootstrap are very similar to the wild bootstrap, we conclude that the specific bootstrapping approach does not make a difference to the results in this application.

Table 2
Effects of QE on output – OLS regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.788**	0.769**	0.721*	0.620	0.526	0.513
	(2.151)	(2.167)	(1.744)	(1.597)	(1.487)	(1.271)
	[0.0314]	[0.0302]	[0.0811]	[0.110]	[0.137]	[0.204]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.072	0.102	0.112	0.043	0.091	0.095
Jarque-Bera	9.567***	6.615**	7.046**	41***	27.17***	25.82***
Skew	0.919	0.770	0.795	1.655	1.395	1.375
Kurtosis	3.761	3.605	3.625	5.510	4.912	4.819
			Outlying Residuals			
Paper ID 1	3	3	3	3	3	3
Std. Res. 1	3.598	3.504	3.473	4.282	4.032	3.943
Paper ID 2				46		
Std. Res. 2				3.028		
Panel B: Standardised Effect						
CB Affiliation	0.164**	0.162**	0.152*	0.140**	0.127**	0.122*
	(2.327)	(2.408)	(1.885)	(2.119)	(2.094)	(1.656)
	[0.0200]	[0.0160]	[0.0594]	[0.0341]	[0.0363]	[0.0977]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.059	0.170	0.207	0.048	0.078	0.106
Jarque-Bera	207.4***	214.2***	169.9***	595.8***	558.7***	461.5***
Skew	2.525	2.468	2.260	3.577	3.488	3.256
Kurtosis	10.77	11.02	10.06	17.13	16.66	15.33
			Outlying Residuals			
Paper ID 1	3	3	3	3	3	3
Std. Res. 1	4.390	4.975	4.846	5.241	5.268	5.096
Paper ID 2	11	11	11	11	11	11
Std. Res. 2	5.355	5.028	4.771	6.618	6.551	6.298

Note: Regression coefficients shown are estimated with OLS. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. The Jarque-Bera, Skew and Kurtosis statistics are all calculated based on the residuals of a single OLS model estimate with standard errors clustered at the paper-level. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. The 5% significance level for the Jarque-Bera statistic is 5.99.

Table 3

Effects of QE on inflation – OLS regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	1.410***	1.494***	1.278***	1.701**	1.688**	1.393*
	(3.421)	(3.319)	(2.714)	(2.161)	(2.157)	(1.868)
	[0.0006]	[0.0009]	[0.0066]	[0.0307]	[0.0310]	[0.0617]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.142	0.239	0.301	0.126	0.126	0.195
Jarque-Bera	92.23***	35.83***	23.30***	202.5***	200***	126.1***
Skew	1.921	1.286	0.985	-0.777	-0.727	-0.632
Kurtosis	8.196	6.099	5.583	12.45	12.40	10.45
			Outlying Residuals			
Paper ID 1	16	16	16	16	16	16
Std. Res. 1	5.410	4.896	4.683	4.163	4.288	3.968
Paper ID 2				47	47	47
Std. Res. 2				-6.132	-6.241	-5.864
Panel B: Standardised Effect						
CB Affiliation	0.195***	0.226***	0.200**	0.203**	0.218**	0.189**
	(2.600)	(2.676)	(2.561)	(2.256)	(2.280)	(2.203)
	[0.0093]	[0.0075]	[0.0104]	[0.0240]	[0.0226]	[0.0276]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.109	0.249	0.298	0.105	0.185	0.224
Jarque-Bera	603.4***	446.4***	401.3***	570.9***	404.7***	333.9***
Skew	3.312	2.805	2.624	2.724	2.404	2.252
Kurtosis	18.14	16.06	15.42	18.13	15.65	14.44
			Outlying Residuals			
Paper ID 1	14	14	14	14	14	14
Std. Res. 1	8.633	8.336	8.173	8.376	7.956	7.629

Note: Regression coefficients shown are estimated with OLS. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. The Jarque-Bera, Skew and Kurtosis statistics are all calculated based on the residuals of a single OLS model estimate with standard errors clustered at the paper-level. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. The 5% significance level for the Jarque-Bera statistic is 5.99.

We also provide calculations of skewness, kurtosis and a Jarque-Bera test for normality of the residuals. These statistics are calculated based on the one set of residuals obtained from the corresponding OLS regression without bootstrapping. The additional statistics reveal that the residuals in all of these regressions have (often very) high levels of kurtosis. As a result, the Jarque-Bera test rejects normality of the distribution of residuals in every single instance in tables 2 and 3[3].

Rejection of the normal distribution assumption for the residuals can often be due a few influential observations, especially when the sample size isn't large. We studentise[4] the residuals to identify residual observations which are more than three standard deviations from the mean. As can be seen in tables 2 and 3, there is at least one such observation in every regression specification. In tables 2A and 3A, we rerun the regression in tables 2 and 3, but without observations associated with a studentised residual greater than 3.

Table 2A shows that without these influential observations, most of the mean estimates fall by 20%-50%, with an average decline in effect of 30% in Panel A and 40% in Panel B. Similarly, in table 3A, the estimates fall between 5% and 40%, with an average decline of 20% in Panel A and 30% in Panel B. But the majority of estimates remain statistically significant. However, kurtosis remains very high in most specifications and the Jarque-Bera test rejects the hypothesis that the residuals are normally distributed, even when these influential observations are removed. Overall, this shows that the residual distribution continues to exhibit much fatter tails and higher skewness relative to a normal distribution, even if a few very influential observations are removed.

[3] The Jarque-Bera statistic is distributed as $\chi^2(2)$.. The 5% significance level 5.99, the 1% level 9.21 and the 0.1% level is 13.82. The values of over 100 shown in tables 2 and 3 have p-values of less than 2×10^{-22}.

[4] By 'studentise' we mean that we subtract the mean from the residuals, and then divide the resulting series by the standard deviation. A studentised value of 3 indicates an observation 3 standard deviations away from the mean residual.

Table 2A
Effects of QE on output – OLS regressions with outliers excluded

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.613*	0.623*	0.559	0.231	0.355	0.325
	(1.825)	(1.847)	(1.456)	(0.792)	(1.137)	(0.928)
	[0.0679]	[0.0648]	[0.145]	[0.428]	[0.255]	[0.354]
Country FE		X	X		X	X
Controls			X			X
Observations	57	57	57	55	56	56
R^2	0.054	0.082	0.096	0.010	0.055	0.058
Jarque-Bera	4.219	2.799	3.236	14.80***	17.82***	19.03***
Skew	0.664	0.537	0.584	1.241	1.276	1.301
Kurtosis	2.886	2.842	3.005	3.543	4.061	4.177
Panel B: Standardised Effect						
CB Affiliation	0.103**	0.106**	0.0974*	0.0713**	0.0642*	0.0615
	(2.045)	(2.095)	(1.761)	(2.043)	(1.832)	(1.560)
	[0.0409]	[0.0361]	[0.0782]	[0.0411]	[0.0670]	[0.119]
Country FE		X	X		X	X
Controls			X			X
Observations	56	56	56	55	55	55
R^2	0.060	0.233	0.248	0.061	0.094	0.097
Jarque-Bera	24.60***	7.286**	5.383*	19.65***	15.60***	16.66***
Skew	1.310	0.807	0.723	1.338	1.204	1.229
Kurtosis	4.918	3.718	3.465	4.188	4.003	4.108

Note: Regression coefficients shown are estimated with OLS with outliers excluded. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. The Jarque-Bera, Skew and Kurtosis statistics are all calculated based on the residuals of a single OLS model estimate with standard errors clustered at the paper-level. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. The 5% significance level for the Jarque-Bera statistic is 5.99.

The leptokurtic nature of the residuals, and the rejection of a normal distribution for each specification, requires an econometric approach that, unlike the OLS estimator, is robust to distributional assumptions about the residuals. A standard approach in the econometrics literature in this case is the median regression estimator. We re-estimate tables 2 and 3 with this estimator in tables 2B and 3B. On average, the mean estimates decline by 50% relative to those shown in tables 2 and 3. Out of 24 regression specifications, only two show statistical significance at the 5% level. Testing at the 5% level, one would expect one of these 24 regressions to be statistically significant at random. Once the control variables are included, no specification is statistically significant.

Table 3A

Effects of QE on inflation – OLS regressions with outliers excluded

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	1.286***	1.346***	1.209***	0.973**	0.948**	0.875*
	(3.244)	(3.187)	(2.762)	(2.325)	(2.169)	(1.951)
	[0.0012]	[0.0014]	[0.0057]	[0.0201]	[0.0301]	[0.0510]
Country FE		X	X		X	X
Controls			X			X
Observations	52	52	52	51	51	51
R^2	0.179	0.230	0.281	0.124	0.128	0.155
Jarque-Bera	10.66***	4.474	2.616	20.28***	19.60***	17.46***
Skew	1.062	0.706	0.526	1.261	1.246	1.148
Kurtosis	3.638	2.733	2.682	4.783	4.735	4.717
Panel B: Standardised Effect						
CB Affiliation	0.132***	0.156***	0.146***	0.137**	0.142**	0.131**
	(3.001)	(3.184)	(2.847)	(2.264)	(2.365)	(2.195)
	[0.0027]	[0.0015]	[0.0044]	[0.024]	[0.0180]	[0.0282]
Country FE		X	X		X	X
Controls			X			X
Observations	52	52	52	52	52	52
R^2	0.119	0.245	0.296	0.111	0.161	0.182
Jarque-Bera	18.92***	4.741*	2.950	80.89***	51.79***	45.13***
Skew	1.266	0.734	0.583	-0.916	-0.813	-0.784
Kurtosis	4.524	3.177	2.954	8.829	7.610	7.286

Note: Regression coefficients shown are estimated with OLS with outliers excluded. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. The Jarque-Bera, Skew and Kurtosis statistics are all calculated based on the residuals of a single OLS model estimate with standard errors clustered at the paper-level. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. The 5% significance level for the Jarque-Bera statistic is 5.99.

Table 2B
Effects of QE on output – Median regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.743	0.450	0.288	0.384	0.413	0.346
	(1.594)	(0.918)	(0.540)	(1.061)	(1.137)	(0.851)
	[0.111]	[0.359]	[0.589]	[0.289]	[0.256]	[0.395]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.072	0.091	0.067	0.043	0.040	0.019
Panel B: Standardised Effect						
CB Affiliation	0.0624	0.0360	0.0408	0.0619	0.0605	0.0302
	(1.002)	(0.576)	(0.535)	(1.333)	(1.588)	(0.672)
	[0.316]	[0.565]	[0.592]	[0.182]	[0.112]	[0.502]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.059	0.149	0.169	0.048	0.054	0.046

Note: Regression coefficients shown are estimated with a median (quantile) regression. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

While median regressions are a common approach to examine the robustness of OLS results, they protect only against vertical outliers. Bad leverage points could still affect the results. MM and MS regressions, the latest evolution of the robust regression method in Huber (1964), are also robust to bad leverage points. We follow the recommendation of Verardi and Croux (2008) and use MS regression when the specification includes dummy variables, which in this case are country fixed effects. Tables 2C and 3C show the results from the analysis with the MM/MS regression. These estimates are on average 75% smaller than the estimates in tables 2 and 3. This is due to a decline of close to 100% in a few cases in table 2C, but a 70% decline can be observed across the board in table 3C. The p-values in both tables show that these results are far away from conventional statistical significance levels.

Table 3B
Effects of QE on inflation – Median regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	1.004**	0.940*	0.570	0.733*	0.813*	0.360
	(2.029)	(1.713)	(1.032)	(1.647)	(1.658)	(0.680)
	[0.0425]	[0.0866]	[0.302]	[0.0996]	[0.0972]	[0.496]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.142	0.131	0.282	0.126	0.120	0.143
Panel B: Standardised Effect						
CB Affiliation	0.102	0.126	0.0685	0.111**	0.0793	0.0749
	(1.643)	(1.618)	(0.877)	(2.126)	(1.506)	(1.420)
	[0.100]	[0.106]	[0.380]	[0.0335]	[0.132]	[0.156]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.109	0.244	0.233	0.105	0.180	0.178

Note: Regression coefficients shown are estimated with a median (quantile) regression. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

One concern with the MM/MS regression approach is that standard errors of these models cannot be clustered at the regression level. However, it is possible to cluster at the bootstrap level. To examine if the lack of regression-level clustering makes a large difference to the results, we re-estimate the OLS (tables 2 and 3) and median (tables 2B and 3B) regression models with only the bootstrap cluster. These results are presented in appendix A2. Comparing these results to tables 2 and 3 for OLS and 2B and 3B for the quantile regressions shows the lack of regression-level clustering doesn't not affect the p-values very much. We therefore conclude that the lack of clustering at regression level has a very small (if any) effect on the p-values generated by the MM/MS regression approach.

Table 2C
Effects of QE on output – MM/MS regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.470	-0.0555	-0.115	0.339	0.289	0.281
	(1.156)	(-0.0695)	(-0.126)	(1.282)	(0.728)	(0.480)
	[0.248]	[0.945]	[0.900]	[0.200]	[0.467]	[0.631]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
Panel B: Standardised Effect						
CB Affiliation	0.0184	0.0159	-4.52e-05	0.0152	0.0299	0.0268
	(0.439)	(0.204)	(-0.000353)	(0.611)	(0.776)	(0.400)
	[0.660]	[0.838]	[1.000]	[0.541]	[0.438]	[0.689]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57

Note: Regression coefficients are estimated with an MM regression in specification (1) and an MS regression in specifications (2) and (3). Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

Table 3C
Effects of QE on inflation – MM/MS regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.571	0.655	0.174	0.350	0.529	0.462
	(1.628)	(1.068)	(0.205)	(0.850)	(0.941)	(0.560)
	[0.103]	[0.285]	[0.837]	[0.395]	[0.347]	[0.576]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
Panel B: Standardised Effect						
CB Affiliation	0.0638	0.0497	0.0439	0.0525	0.0334	0.0638
	(1.400)	(0.597)	(0.364)	(1.237)	(0.553)	(0.724)
	[0.161]	[0.551]	[0.716]	[0.216]	[0.581]	[0.469]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53

Note: Regression coefficients are estimated with an MM regression in specification (1) and an MS regression in specifications (2) and (3). Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

Fabo et al (2021) also run limited dependent variable regressions to assess whether central bank researchers are more likely to report statistical and economic significance of their results (Table 4). A similar approach is used to study whether central bank researchers who wrote QE papers had accelerated career progression (Table 6). The robust regression models employed in this paper are not valid when the dependent variable takes only a small number of discrete values. However, we can still examine whether the residuals are consistent with assumption of a normal distribution. While the value of kurtosis is closer to 3 in some of the specifications in Tables 4A and 6, the Jarque-Bera statistics support rejection of the normality assumption in most specification as the degree of skewness is inconsistent with a normal distribution. Moreover, unless the product of the true parameter estimate and each observation lies between zero and one, OLS estimates of linear probability models will be biased and inconsistent[5]. OLS inference in this application might be affected by the presence of outliers in these regressions as well as concerns about inconsistency in the estimates.

While it might be thought that significance was a dichotomous variable, Fabo *et al.* (2021) identify some ambiguous cases. Clearly insignificant results are given a significance of 0; those that are clearly significant are given a value of 1, and the ambiguous cases are given significance of 0.5. When the explanatory power of the linear model is small, as is the case particularly for inflation significance, the residuals of the model will not be very different from the dependent variables and since these take one of only three values it is most unlikely that the residuals will be normally distributed. An ordered logit model, in contrast, is designed to model a situation where there is only a small number of ranked outcomes.

[5] See Horrace and Oaxaca (2006) for a discussion.

Table 4A
Statistical and Economic Significance- Linear Model

	Statistical Significance			Economic Significance		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect on Output						
CB Affiliation	0.412**	0.388**	0.367**	0.335***	0.344***	0.399***
	(2.553)	(2.528)	(2.236)	(2.780)	(2.795)	(3.285)
	[0.011]	[0.012]	[0.025]	[0.005]	[0.005]	[0.001]
Country FE		X	X		X	X
Controls			X			X
Observations	41	41	41	66	66	66
Identifier 1						12
Std. Res. 1						-3.104
R^2	0.233	0.280	0.295	0.139	0.145	0.250
Jarque-Bera	13.33***	11.39***	11.20***	16.62***	16.66***	11.71***
Skewness	-1.304	-1.194	-1.212	-1.200	-1.197	-1.003
Kurtosis	4.002	3.983	3.827	3.533	3.572	3.487

	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: Effect on Inflation						
CB Affiliation	0.202	0.202	0.163	0.196	0.207	0.248*
	(1.246)	(1.275)	(1.045)	(1.314)	(1.369)	(1.772)
	[0.213]	[0.202]	[0.296]	[0.189]	[0.171]	[0.076]
Country FE		X	X		X	X
Controls			X			X
Observations	38	38	38	60	60	60
R-squared	0.044	0.118	0.206	0.041	0.043	0.137
Jarque-Bera	21.54***	17.91***	10.95***	13.52***	13.08***	10.85***
Skewness	-1.727	-1.554	-1.254	-1.159	-1.139	-1.041
Kurtosis	4.297	4.285	3.792	2.825	2.789	3.053

Note: Panel A show results for a linear probability model where researchers report either economic or statistical significance of the output effect of QE. while panel B shows results for the inflation effect of QE. Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. The Jarque-Bera, Skew and Kurtosis statistics are all calculated based on the residuals of a single OLS model estimate with standard errors clustered at the paper-level. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. The 5% significance level for the Jarque-Bera statistic is 5.99.

Importantly, logit models are also an econometrically consistent way of modelling these data. In table 4B we show the results estimated using one of these. This cannot be estimated when examining the statistical significance of the effect on output, because some observations are completely determined by the CB Affiliation variable. The results in table 4B show significance of central bank affiliation in only one of the specifications for economic significance of the output effect. But now the economic significance of the inflation effect is significant in two specifications, one more than in table 4A. Although, there is one less significant result with the logit model in 4b, this evidence is clearly less convincing than the evidence for tables 2 and 3.

Table 4B
Statistical and Economic Significance- Logit Model

	Statistical Significance			Economic Significance		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect on Output						
CB Affiliation				2.480	2.522***	4.250
				(0.711)	(2.889)	(1.578)
				[0.477]	[0.00386]	[0.115]
Country FE					X	X
Controls						X
Observations				66	66	66
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: Effect on Inflation						
CB Affiliation	1.591	1.558	2.682	0.985	1.206*	1.749**
	(0.177)	(0.0689)	(0)	(1.010)	(1.699)	(2.278)
	[0.859]	[0.945]	[1]	[0.313]	[0.0893]	[0.0227]
Country FE		X	X		X	X
Controls			X			X
Observations	38	38	38	60	60	60

Note: This table shows the results of estimating ordered logit models. The z-statistics and p-values are estimated using 10,000 replications with the pairs cluster bootstrap. When there are no country fixed effects the estimation equations are also clustered by paper. z-statistics are reported in (), and p-values in []. The model cannot be estimated in the first three cases because some observations are fully explained. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

Fabo *et al.* (2021) also collect data on the tone of the papers' abstracts. Specifically, they compute the share of positive and the share of negative adjectives describing the results in the abstract of each paper. They refer to the difference of the positive share with the negative share as the 'sentiment' score of each paper. Table 5 replicates these results in our paper. Table 5 shows that the results are robust regardless of econometric estimator. This is should not be surprising because Jarque-Bera statistics of less than 6 in table 5 do not allow us to reject the null-hypothesis of a Gaussian distribution.

Table 5
Tone of the Abstract (Sentiment score) regressions

	(1)	(2)	(3)
Panel A: OLS regression			
CB Affiliation	0.0464**	0.0532***	0.0556**
	(2.075)	(2.587)	(2.561)
	[0.0379]	[0.00967]	[0.0104]
Observations	54	54	54
R^2	0.081	0.129	0.133
Jarque-Bera	3.988	3.989	3.243
Skewness	0.567	0.402	0.350
Kurtosis	3.698	4.061	3.976
Identifier 1	51	51	51
Std. Res. 1	3.257	3.421	3.398
Panel B: OLS regression with outliers excluded			
CB Affiliation	0.0370*	0.0453**	0.0461**
	(1.777)	(2.343)	(2.275)
	[0.0755]	[0.0191]	[0.0229]
Observations	53	53	53
R^2	0.062	0.137	0.146
Jarque-Bera	0.409	0.508	1.018
Skewness	0.213	-0.222	-0.331
Kurtosis	3.064	3.180	3.152
Panel C: Median regressions			
CB Affiliation	0.0710***	0.0760***	0.0676**
	(4.374)	(3.062)	(2.523)
	[1.22e-05]	[0.00220]	[0.0116]
Observations	54	54	54
R^2	0.081	0.117	0.115
Panel D: MM/MS regressions			
CB Affiliation	0.0738***	0.0787**	0.0762*
	(3.397)	(2.083)	(1.703)
	[0.000682]	[0.0372]	[0.0886]
Country FE		X	X
Controls			X

Note: This table shows coefficients estimated with different regression methods. Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. The Jarque-Bera, Skweness and Kurtosis statistics reported in panels A and B are obtained from corresponding OLS regressions with clustered standard errors only. Control variables include the number of authors and the log of three plus average author experience. t-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. The 5% significance level for the Jarque-Bera statistic is 5.99.

The values of the Jarque-Bera statistic in table 6 suggest that most of the specifications of the equations explaining career outcome may be affected by non-Gaussian residuals. But it is important to note that the degree of statistical significance based on the pairwise bootstrap is much smaller relative to the wild bootstrap shown in table A_6. Logit models are still biased in small samples. Finally, the results remained unchanged when we excluded identified outliers. All of these challenges suggest that it is hard to find evidence that lack of Gaussianity is a problem.

Table 6
Career Outcomes and the Effects of QE on Output

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total Programme Effect						
Output Effect	0.264**	0.219	0.485*	0.204	0.204	0.460
	(1.985)	(1.458)	(1.655)	(1.455)	(1.131)	(1.169)
	[0.047]	[0.145]	[0.098]	[0.146]	[0.258]	[0.242]
Country FE		X	X		X	X
Controls			X			X
Observations	34	34	31	32	32	30
R^2	0.030	0.066	0.553	0.027	0.076	0.550
Jarque-Bera	7.490**	6.174**	2.153	14.50***	14.90***	14.08***
Skewness	-0.675	-0.392	0.618	-0.599	-0.358	1.245
Kurtosis	4.862	4.935	3.372	6.072	6.265	5.251
Identifier 1	27	27		27	27	15
Std. Res. 1	-3.243	-3.340		-3.675	-3.806	4.031
Panel B: Standardised Effect						
Output Effect	1.407	1.009	2.661	2.311*	1.838	4.095
	(1.403)	(1.086)	(1.419)	(1.849)	(1.310)	(1.400)
	[0.161]	[0.277]	[0.156]	[0.064]	[0.190]	[0.161]
Country FE		X	X		X	X
Controls			X			X
Observations	34	34	31	32	32	30
R^2	0.044	0.062	0.553	0.051	0.081	0.569
Jarque-Bera	6.120**	6.208**	1.139	13.98***	15.22***	13.42***
Skewness	-0.598	-0.437	0.470	-0.512	-0.381	1.243
Kurtosis	4.699	4.902	2.989	6.071	6.292	5.134
Identifier 1	27	27		27	27	15
Std. Res. 1	-3.188	-3.359		-3.646	-3.807	3.988

Note: Panel A show results for a limited dependent variable OLS model where researchers career progress is a function of their report QE output effects for the total effect of the QE programme. Panel B shows results for the standardised effect. Standard errors are clustered at the paper-level, while p-values are obtained with 10,000 replications of the pairs cluster bootstrap. Control variables include the number of authors and the log of three plus average author experience. The Jarque-Bera, Skew and Kurtosis statistics are all calculated based on the residuals of a single OLS model estimate with standard errors clustered at the paper-level. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. The 5% significance level for the Jarque-Bera statistic is 5.99.

4 Conclusion

In response to the 2008/2009 Global Financial Crisis, central banks introduced QE. Central bank and academic researchers have published several papers on the effect of this new monetary policy tool on output and inflation. Recent work by Fabo *et al.* (2021) provided statistical evidence that central bank researchers systematically reported larger effects of this policy than academic researchers. Central bankers are also most likely to report results as economically/statistically significant, experience career advancement and describe their findings in a more positive light. This work has been influential, and was, for example, referred to the UK's House of Lords report (House of Lords, 2021) and the Bank of England Independent Evaluation Office report on the Bank of England's QE.

While Fabo *et al.* (2021) show that their results are robust to a large number of econometric perturbations, they rely on OLS estimators throughout. But OLS estimates can be subject to outliers. Vertical outliers and bad leverage points can affect OLS estimates, especially when the distribution of the residuals is not Gaussian. The values of skewness and kurtosis in Fabo *et al.* (2021) lead to rejection of the null hypothesis that regression residuals follow a Gaussian distribution in most specifications. This means that their results could be biased by the presence of outliers. Where this is feasible, we re-estimate their results with two econometric methods that are robust to this issue, the median (least absolute deviation) and MM/MS regression. Once these methods are used, the null hypothesis of equality in effects among central bank and academic researchers cannot be rejected any more. We don't find strong evidence that the lack of Gaussian residuals is an issue in the other regressions.

Fabo *et al.* (2021) have started an interesting research agenda, comparing the research results which originate in academia and central banks. This is clearly a very important task.

However, given the importance of such results in the public debate, it is imperative that they emerge as a consensus among different studies with different econometric approaches. We show that the null hypothesis of a Gaussian distribution of the regression residuals can be rejected in the majority of their specifications. General readers are likely to be unaware of the possible pitfalls of using the OLS estimator when this is the case. We show that some of their results that central bankers systematically report higher QE multipliers are not sustained when subject to standard robust regression methods. At the same time, there is no evidence that their other results are affected by the lack of a Gaussian distribution in the residuals. Whether or not central bank research systematically reaches different conclusions than academic research therefore remains an important topic for future research.

5 References

Bertrand, M., E. Duflo, and S. Mullainathan (2004). "How much should we trust differences-in-differences estimates?". *Quarterly Journal of Economics.* Vol. 119. Pp 249-275.

Edgeworth, F. (1887). "On observations relating to several quantities". *Hermathena*. Vol 6. Pp 279-285.

Fabo, B., M. Jančokova, E. Kempf and L. Pástor (2021) "Fifty shades of QE: comparing findings of central bankers and academics". *Journal of Monetary Economics*. Vol. 120. Pp. 1-20. https://doi.org/10.1016/j.jmoneco.2021.04.001

Huber, P. (1964). "Robust estimation of a location parameter". *Annals of Mathematical Statistics.* Vol. 35. Pp. 73-101. https://www.jstor.org/stable/2238020.

Horrace, W. C. and R. L. Oaxaca (2006) "Results on the bias and inconsistency of ordinary least squares for the linear probability model". *Economics Letters*, Vol. 90, Pp. 321-327.

Koenker, R. and G. Bassett. (1978). "Regression quantiles". *Econometrica*. Vol 46. Pp. 33-50.

Maronna, R. and V. Yohai. (2000). "Robust regression with both continuous and categorical predictors" *Journal of Statistical Planning and Inference*. Vol. 89. Pp 197-214.

MacKinnon, J.G. and M.D. Webb (2017). "Pitfalls when estimating treatment effects with clustered data". *The Political Methodologist*. Vol. 24. Pages 20-31.

Rousseeuw P and A. Leroy. (2003). *Robust regression and outlier detection*. Wiley. Hoboken.

Veradi, V. and C. Croux. (2009). "Robust regression in Stata". *Stata Journal*. https://doi.org/10.1177/1536867X0900900306

Weale M. and T. Wieladek. (2016). "What are the macroeconomic effects of asset purchases?". *Journal of Monetary Economics*. Vol. 79. Pp. 81-93.

Wu, C. F. J.(1986). "Jackknife, bootstrap and other resampling methods in regression analysis". *Annals of Statistics.* Vol. 14. Pp. 1261-1295.

Yohai, V. (1987). "High breakdown-point and high efficiency robust estimates for regression". *Annals of Mathematical Statistics*. Vol. 15. Pp 642-656. https://www.jstor.org/stable/2241331

APPENDIX A1 – Regression tables with the wild bootstrap

Table A_2

Effects of QE on output – OLS regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.788**	0.769**	0.721*	0.620	0.526	0.513
	(2.152)	(2.164)	(1.829)	(1.604)	(1.513)	(1.366)
	[0.0366]	[0.0386]	[0.0901]	[0.117]	[0.129]	[0.179]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.072	0.102	0.112	0.043	0.091	0.095
Panel B: Standardised Effect						
CB Affiliation	0.164**	0.162**	0.152*	0.140**	0.127**	0.122**
	(2.368)	(2.467)	(2.098)	(2.170)	(2.165)	(1.904)
	[0.0211]	[0.0149]	[0.0521]	[0.0216]	[0.0209]	[0.0480]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.059	0.170	0.207	0.048	0.078	0.106

Note: Regression coefficients shown are estimated with OLS. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level and wild bootstrap clustered at the paper-level. Control variables include the number of authors and the log of three plus average author experience. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. T-statistics reported in (). P-values reported in [].

Table A_3
Effects of QE on inflation – OLS regressions

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	1.410***	1.494***	1.278**	1.701***	1.688**	1.393**
	(3.424)	(3.330)	(2.804)	(2.243)	(2.199)	(2.051)
	[0.00150]	[0.00280]	[0.0117]	[0.00890]	[0.0123]	[0.0438]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.142	0.239	0.301	0.126	0.126	0.195
Panel B: Standardised Effect						
CB Affiliation	0.195***	0.226***	0.200***	0.203***	0.218***	0.189***
	(2.584)	(2.710)	(2.732)	(2.287)	(2.295)	(2.404)
	[0.00590]	[0.00230]	[0.00400]	[0.00380]	[0.00510]	[0.00680]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.109	0.249	0.298	0.105	0.185	0.224

Note: Regression coefficients shown are estimated with OLS. Panel A shows results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are clustered at the paper-level and wild bootstrap clustered at the paper-level. Control variables include the number of authors and the log of three plus average author experience. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. T-statistics reported in (). P-values reported in [].

Table A_4

Significance

	Statistical Significance			Economic Significance		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect on Output						
CB Affiliation	0.412**	0.388**	0.367**	0.335***	0.344***	0.399***
	(2.417)	(2.409)	(2.193)	(2.777)	(2.780)	(3.423)
	[0.045]	[0.032]	[0.050]	[0.017]	[0.016]	[0.005]
Country FE		X	X		X	X
Controls			X			X
Observations	41	41	41	66	66	66
R^2	0.233	0.280	0.295	0.139	0.145	0.250
Panel B: Effect on Inflation						
CB Affiliation	0.202	0.202	0.163	0.196	0.207	0.248*
	(1.182)	(1.249)	(1.089)	(1.290)	(1.364)	(1.865)
	[0.341]	[0.291]	[0.378]	[0.219]	[0.200]	[0.083]
Country FE		X	X		X	X
Controls			X			X
Observations	38	38	38	60	60	60
R^2	0.044	0.118	0.206	0.041	0.043	0.137

Note: Regression coefficients shown are estimated with OLS. Control variables include the number of authors and the log of three plus average author experience. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. Standard errors are clustered at the paper-level and wild bootstrap clustered at the paper-level. t-statistics reported in (). P-values reported in [].

Table A_5

Tone of Abstract (Sentiment score) regressions

	Peak Effect		
	(1)	(2)	(3)
Panel A: OLS regression			
CB Affiliation	0.0464**	0.0532**	0.0556**
	(2.052)	(2.592)	(2.597)
	[0.0541]	[0.0134]	[0.0127]
Observations	54	54	54
R^2	0.081	0.129	0.133
Country FE		X	X
Controls			X

Note: Regression coefficients shown are estimated with OLS. Control variables include the number of authors and the log of three plus average author experience. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. Standard errors are clustered at the paper-level and wild bootstrap clustered at the paper-level. t-statistics reported in (). P-values reported in [].

Table A_6

Career Outcomes and the Effects of QE on Output

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total Programme Effect						
Output Effect	0.264**	0.219**	0.485**	0.204*	0.204	0.460**
	(2.324)	(1.849)	(2.646)	(1.784)	(1.254)	(2.124)
	[0.0299]	[0.0368]	[0.0181]	[0.0762]	[0.231]	[0.0192]
Country FE		X	X		X	X
Controls			X			X
Observations	34	34	31	32	32	30
R^2	0.030	0.066	0.553	0.027	0.076	0.550
Panel B: Standardised Effect						
Output Effect	1.407	1.009	2.661*	2.311**	1.838	4.095**
	(1.407)	(1.149)	(1.857)	(2.001)	(1.454)	(2.147)
	[0.228]	[0.354]	[0.0866]	[0.0419]	[0.131]	[0.0185]
Country FE		X	X		X	X
Controls			X			X
Observations	34	34	31	32	32	30
R^2	0.044	0.062	0.553	0.051	0.081	0.569

Note: Regression coefficients shown are estimated with OLS. Control variables include the number of authors and the log of three plus average author experience. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1. Standard errors are clustered at the paper-level and wild bootstrap clustered at the paper-level. t-statistics reported in (). P-values reported in [].

APPENDIX A2 – Regression tables with cluster bootstrap only

Table A2_2

Effects of QE on output – OLS regressions with cluster bootstrap only

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.788**	0.769**	0.721*	0.620	0.526	0.513
	(2.143)	(2.146)	(1.734)	(1.611)	(1.476)	(1.270)
	[0.0321]	[0.0319]	[0.0828]	[0.107]	[0.140]	[0.204]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.072	0.102	0.112	0.043	0.091	0.095
Panel B: Standardised Effect						
CB Affiliation	0.164**	0.162**	0.152*	0.140**	0.127**	0.122*
	(2.325)	(2.438)	(1.887)	(2.142)	(2.081)	(1.687)
	[0.0201]	[0.0147]	[0.0591]	[0.0322]	[0.0374]	[0.0917]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.059	0.170	0.207	0.048	0.078	0.106

Note: Regression coefficients shown are estimated with OLS . Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are only cluster bootstrapped at the paper level. Control variables include the number of authors and the log of three plus average author experience. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

Table A2_3

Effects of QE on inflation – OLS regressions with cluster bootstrap only

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	1.410***	1.494***	1.278***	1.701**	1.688**	1.393*
	(3.440)	(3.279)	(2.668)	(2.187)	(2.183)	(1.901)
	[0.000582]	[0.00104]	[0.00762]	[0.0288]	[0.0290]	[0.0573]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.142	0.239	0.301	0.126	0.126	0.195
Panel B: Standardised Effect						
CB Affiliation	0.195**	0.226***	0.200**	0.203**	0.218**	0.189**
	(2.558)	(2.704)	(2.552)	(2.264)	(2.324)	(2.216)
	[0.0105]	[0.00685]	[0.0107]	[0.0235]	[0.0201]	[0.0267]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.109	0.249	0.298	0.105	0.185	0.224

Note: Regression coefficients shown are estimated with OLS with outliers excluded. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are only cluster bootstrapped at the paper level. Control variables include the number of authors and the log of three plus average author experience. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

Table A2_2B

Effects of QE on output – Median regressions with cluster bootstrap only

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	0.743	0.450	0.288	0.384	0.413	0.346
	(1.573)	(0.915)	(0.540)	(1.044)	(1.123)	(0.840)
	[0.116]	[0.360]	[0.589]	[0.297]	[0.261]	[0.401]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.072	0.091	0.067	0.043	0.040	0.019
Panel B: Standardised Effect						
CB Affiliation	0.0624	0.0360	0.0408	0.0619	0.0605	0.0302
	(0.978)	(0.569)	(0.526)	(1.321)	(1.618)	(0.667)
	[0.328]	[0.569]	[0.599]	[0.186]	[0.106]	[0.505]
Country FE		X	X		X	X
Controls			X			X
Observations	58	58	58	57	57	57
R^2	0.059	0.149	0.169	0.048	0.054	0.046

Note: Regression coefficients shown are estimated with a median (quantile) regression. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are only cluster bootstrapped at the paper level. Control variables include the number of authors and the log of three plus average author experience. T-statistics reported in (). P-values reported in []. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.

Table A2_3B
Effects of QE on inflation – Median regressions with cluster bootstrap only

	Peak Effect			Cumulative Effect		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Total Program Effect						
CB Affiliation	1.004**	0.940*	0.570	0.733	0.813*	0.360
	(2.017)	(1.747)	(1.041)	(1.611)	(1.664)	(0.671)
	[0.0437]	[0.0807]	[0.298]	[0.107]	[0.0962]	[0.502]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.142	0.131	0.282	0.126	0.120	0.143
Panel B: Standardised Effect						
CB Affiliation	0.102*	0.126	0.0685	0.111**	0.0793	0.0749
	(1.649)	(1.575)	(0.877)	(2.126)	(1.498)	(1.399)
	[0.0992]	[0.115]	[0.380]	[0.0335]	[0.134]	[0.162]
Country FE		X	X		X	X
Controls			X			X
Observations	53	53	53	53	53	53
R^2	0.109	0.244	0.233	0.105	0.180	0.178

Note: Regression coefficients shown are estimated with a median (quantile) regression. Panel A show results for the total QE program effect, while panel B shows results which are standardised to a 1% rise in QE (as a share of GDP). Standard errors are only cluster bootstrapped at the paper level. Control variables include the number of authors and the log of three plus average author experience. *** if p-value < .01, ** if p-value < .05 and * if p-value < .1.