Dear PSRM Editors and Reviewers

Thank you for the opportunity to revise and resubmit our manuscript: "Inflated Expectations: How government partisanship shapes monetary policy bureaucrats' inflation forecasts" as well as your very helpful suggestions.

We have incorporated the reviewers' suggestions in the revised manuscript and provide discussions of each change below.

It is important to note that in all of the alternative model specifications suggested by the reviewers, the key presidential party ID variable's estimated coefficients remained substantively the same as before in terms of direction, magnitude, and statistical significance. These correspond to the predictions of our presidential partisan heuristics approach, but not the major alternatives. We have significantly improved the discussion of our theory in the revised version of the paper as suggested.

Full bibliographic details for citations in this letter are available in the paper.

Thank you again for your suggestions and we look forward to your comments.

Kind regards,
Christopher Gandrud and Cassandra Grafström

Concerns Shared by Both Reviewers

Added Supplemental Materials Section

To be able to adequately address the reviewers' suggestions within the journal's space constraints we have placed a number of robustness checks in a new Supplementary Materials section at the end of the paper. If published, this section could be made available separately online. Items in the Supplementary Materials include additional models with variables that the

reviewers suggested—past inflation, oil prices, productivity changes, wars, and so on--as well as a number of robustness checks from the original body of the text. In particular, we moved the discussion and analysis of models with matched and orthogonal data to the Supplementary Materials. We discuss changes to these sections suggested by reviewers later in this letter.

Further Developed Theoretical Argument for Presidential Partisan Heuristics

Both reviewers felt that the original theoretical argument was underdeveloped. To address this concern we have significantly improved the discussion of our presidential partisan heuristics argument (see pages 5-9). We have made significant improvements to the overall theory section (section 2) as well and urge the reviewers to re-read this section.

Specific comments we responded to include:

R2: "The starting point for a theory of bias in inflation forecasts should be a comprehensive and clear discussion of the incentives/preferences of staff (there is some ad-hoc discussion of that on p. 6). What do they aim for when forecasting inflation? Who is their principal? What happens when they are wrong? What makes them want to get it "right"? What makes it likely that staff will not update?"

- Our original discussion of Fed Staff forecasters' motivation was unclear. To address this we explicitly start with the assumption that Fed Staff actually want to create a forecast that is as accurate as possible (p. 5-6). This is plausible because they could both have a preference (because of professional socialisation) and/or an incentive (because of performance reviews and regular private academic analysis that could provoke FOMC or even Congressional scrutiny) to be accurate. Though we don't have direct evidence for some of these mechanisms, multiple pieces of research have found that Greenbook inflation forecasts are generally more accurate than private sector forecasts and other government forecasts.

- Our discussion on why staff may not update is on pages 8-9.

Reviewer 2 continues their critique of our theoretical mechanism and its empirical implications:

R2: "I would like to understand a lot better how "heuristics" come to life and the literature behind them . . . I think authors need to specify their approach better before going into alternative theories. It seems to me that the alternatives need to be considered in a robustness section after the key results."

Likewise, Reviewer 1 found our theoretical exposition inadequate, stating:

R1: "The authors' exposition of three potential causal explanations (partisan heuristics, partisan preferences, and monetary expectations) is clear. And as they show in figure 1, these mechanisms are differentiated in their effects over a presidential term. This should be more clearly articulated in the written text."

- We agree that a more thorough explanation of our causal mechanism and the literature behind it would greatly improve the paper. As such, we expanded the theory by discussing previous research on heuristics generally (p. 6 & 8-9) and related work on the use of rules of thumb among monetary policy-makers (p. 5-6) and how presidential partisanship has been found in numerous pieces of previous research to influence inflation expectations, even among high-information market actors (p. 7). We combine these developed literatures to create our theory of how Fed Staff that want to create accurate forecasts, but face forecasting uncertainty could rely on presidential partisan heuristic intuitions that ultimately result in forecasting errors (p. 7-9).

- We have also clarified in the discussion that the partisan heuristics theory does not assume that Fed Staff have a particular partisan or even an inflation preference. Instead the theory makes the argument that they incorrectly intuit the effect presidents from different partisan backgrounds have on inflation.

- Again, in the original version of the paper (and in this version as well, see p. 8-9) we discuss why a heuristic may not be updated. Primarily, the extensive literature on heuristics shows that they are not 'conscious' judgments. As such, and as we document as being a major motivation of the paper, they are not being explicitly examined and therefore have not been adequately updated. If they were being examined and updated, then they wouldn't be heuristics, but would be rational expectations. The empirical consequence of them being regularly examined and updated is that presidential partisan errors would decrease or even disappear over time, something we do not observe.

Beliefs of Fed Staff that Presidents' Policies Differ by Their Party ID

R1: "The authors make the assumption that forecasters tend to hold the view that Democratic and Republican administrations differ with respect to fiscal policy. Can the authors provide any evidence (e.g., cite a study) corroborating this assumption? There should be some studies that show that economists tend to be more "conservative" when it comes to fiscal policy."

R2: "It seems very broad to claim that Democrats and Republicans "behave differently in government". Is it about a set of policies? If so which ones?"

- We have enhanced the discussion of which economic policies are assumed to differ under Republican and Democratic presidents. We do not have direct evidence that Fed Staff believe that presidents of different political parties are different with respect to fiscal policy. However, to substantiate the plausibility of such a claim we included a discussion of a broad set of research showing that people from a variety of backgrounds (academics, market actors, voters) expect that Democrat and Republican presidents will pursue different economic policies and as such inflation is expected to be higher under Democrats. Please see in particular pages 6-8 for details.

- Furthermore, it's important to reiterate that the heuristic theory is based not on the idea that Fed Staff are 'conservative', just that they incorrectly intuit the effect of presidents from different partisan backgrounds have on inflation. As mentioned earlier, we have clarify our assumptions about staff's preferences in the paper.

Relationship Between Staff and FOMC

Both reviewers posed questions about the relationship between the Staff and the Board of Governors. Both made related arguments about the potential for the Board of Governors' to directly or indirectly affect forecast errors in ways that could produce similar empirical patterns to those predicted by our partisan heuristic theory. For instance:

R2: "Can [the failure of the Fed Staff to update] be related to the Republican preferences of FOMC members rather than a heuristic explanation? Staff need not have a partisan preference, but persistence of bias (and heuristic) may be explained by FOMC member preference." . . . "The authors give the impression that the Greenbook consensus forecast is made independently of what the Fed Governors might think/want. Is this true? Is the consensus forecast truly the predictions (coming from the econometric model and professional opinions) of Fed forecasters only? It is quite plausible that staff forecasts strive to align with the views of the Board of Governors (and their political biases with respect to appropriate monetary policy)."

In a similar vein:

R1: "It is also plausible that the composition of the Board of Governors influences the variation in Greenbook forecasts? Governors are nominated by the President, so there is likely to be some partisan bias of Fed Governors and this may in turn shape actual Fed policy and/or the consensus forecast.

Moreover, it is not uncommon for Fed Staff to generate forecasts with Fed Governors (and to speak with them in general). Thus, it is plausible that the consensus forecast may be unduly influenced by the partisan preferences of Fed Governors."

- First it's important to note that the alternative theories we posit do propose interactions between Staff forecasts and the FOMC. They challenge the assumption that Fed Staff forecasts are not driven by partisan and/or inflation preferences and propose that they act on these preferences in a game with FOMC policy-makers. In the partisan preference theory they have both inflation and partisan preferences (e.g. low inflation preferences that lead them to also have a Republican president preference). In this approach the FOMC effectively shares at least the low inflation preference. In the monetary expectations theory the Staff have a strong price stability preference, but no partisan preference, while the FOMC has a strong partisan preference for Republicans that overrides a constant concern for stable prices. Staff then purposefully bias their forecast to counteract the FOMC's partisan favouring actions. We clarify this in the text in the introduction to section 2 and in more detail in section 2.2. Deciding which approach is more plausible is then a focus of our empirical investigations.

- A key observational issue for any plausible partisan preference theory (either that the FOMC or Staff or both have such a preference) is that we would expect the errors to change in the lead up to elections. Reviewer 2 posits that the persistence of the heuristic *over presidential terms* may be due to a partisan preference rather than a heuristic. We feel this is implausible. There are two reasons why it is unlikely that Fed Staff (or FOMC members) would have a pure partisan preference, but totally disregard inflation. The main partisan preference story regarding central bankers (see Arel-Burdock and Clark 2013) is that they prefer right-leaning governments largely *because* right-leaning governments are more concerned with price stability. They would only want to drop or raise interest rates *before elections* rather than across an entire presidential term, which would risk, for example, high inflation for four years. Also even if they had a pure partisan preference and didn't care about

inflation, it would be electorally counterproductive to, for example, keep interest rates very low for an entire presidential term as it may cause high and persistent inflation thus hurting the incumbent's reelection chances.

- We find no evidence that partisan errors actually change in the run up to elections.

- We have expanded the discussion of how the Fed Staff relate to the FOMC to improve the reader's understanding of the likely direction of influence between the staff and FOMC (p. 12). Primarily we show that Greenbook forecasts are presented before FOMC meetings and that the FOMC spends a considerable amount of time discussing them. We added a discussion supporting the assumption that Greenbook forecasts are largely not just a reflection of FOMC preferences and beliefs, as previous research has shown that the Greenbook and FOMC forecasts are substantively different (see Romer and Romer 2008).

- In regards to Review 1's concern that the partisan preferences of the FOMC change over time due to changes in the partisanship of the appointing president: This is certainly a possibility, though in the time period we examine we wonder what the magnitude and direction of such changes would be. For example, two of the three FOMC chairman appointed or renominated by Democrats--Volker and Greenspan--were hardly strong pro-Democratic partisans willing to use loose monetary policy to aid Democrat's reelection changes (Greenspan of course was a Republican). William McChesney Martin, Jr. was chair for a very short time in our sample During this period inflation was underestimated, the opposite of what a strong partisanship theory would predict, despite there being a Republican president (Nixon). At least at the chair level in the vast majority of our observation period it is reasonable to assume that the FOMC generally had a price stability and/or Republican partisan preference. We do not discuss this in the text, but would be happy to if the reviewers think it is necessary.

Furthermore:

R2: "Under what kind of interaction between staff and FOMC would staff ever want to systematically counter FOMC

policies?..."

- The monetary policy alternative hypothesis directly addresses this issue. It posits a situation where Fed staff have a price stability preference and no partisan preference, but the FOMC is willing to have a little inflation before elections to elect a Republican. The staff then bias their forecasts to counter the FOMC's partisan actions. We did not find evidence for this theory.

- Building on what we discussed above, the other possibilities seem implausible and we do not discuss them in the text. For example, the possibility that the staff or FOMC have a Democratic presidential preference has not been discussed in the literature and, as we mentioned above, seems implausible for at least the FOMC chairs in our time period. Furthermore, the results wouldn't support such stories. Creating information that would justify having high interest rates during Democratic presidencies (the pattern of the bias that we found) seems like a counterproductive way to win elections, especially given what the electoral business cycle literature argues. We know of no previous work positing that either the staff or FOMC over the period we examine have a high inflation preference.

- We would be happy to develop our discussion of this in the paper if requested.

Matched Data and Alternative Parametric Models

There appears to have been some confusion about why we presented results from parametric models with both data that was preprocessed using matching and non-preprocessed data.

R1: "The authors never provide any rationale for using matching techniques. Why is this preferred to a well specified OLS model? Does matching reduce bias – if so, what kind and in what direction?"

R2: "It is not clear what [matching] buys, or what are the

counters to using it. What are the problems in the data that are being solved by matching?"

- Following Ho et al. (2007) our intention was largely to use preprocessed matched data in parametric models as a robustness check of the possibility that our results from models with non-matched data were model dependent, rather than reflecting a real causal effect. In none of the models with matched data did the results substantively change (though the standard errors increased, likely as the result of a large fall in the sample size as noted by Reviewer 1) from models with non-matched data. This bolsters our claim that incorrect expectations about the effects of presidential partisanship play an important role in causing inflation forecasting errors.

- To help emphasise the matching-as-robustness check reasoning we have moved our discussion of this method and the results to the Supplementary Materials section. We also expanded this discussion to better explain why matching is a useful robustness check.

R1: "If authors use matched data to justify a causal relationship between Presidential partisan identification and forecast errors, the paper does not argue this."

- We have also expanded the discussion of the reasons for using matching to help identify causal effects within the Neyman-Rubin framework (see Sekhon 2008) in the Supplementary Materials. Please see that section for details.

In a similar vein:

R2: "The authors should directly explain how the chosen estimation (including the Bayesian normal linear regression) are a good choice for their data, instead of referring the reader to the references."

- The Bayesian normal linear regression model was primarily also a robustness check of the basic normal linear and OLS models. Imai, King, and Lau (2008) argue that Bayesian normal linear regression can

produce more valid inferences for small sample sizes (our sample contains about 135 quarters). That being said, the results across the Bayesian, normal linear, and OLS regressions were substantively similar using both matched and non-matched data (see figures 4 and 10). This further increases our confidence in the robustness of our results. We have clarified this point in the text (p. 16-17). We would be happy to add a longer discussion of the statistical issues behind the models if the Reviewer feels this is necessary.

Composition of Congress Discussion

R1: "In the empirical section, the authors aptly control for the partisan composition of the Congress. However, the composition of Congress (e.g., divided government) is likely to affect the President's capacity to pass fiscal policy. This should be discussed and incorporated into the paper's theory section."

R2: "In the variables section I find protracted the discussion of … Congress party ID – again, this seems appropriate for robustness checks."

- Reviewers 1 and 2 have conflicting advice on how to present the composition of Congress discussion. One suggests moving it to the robustness checks. The other suggests moving it into the main theory section. Given this conflicting advice we have left the discussion in the Variables section of the main text. However, we would be pleased to move it to another section if a consensus is reached.

Reviewer Specific Comments:

Reviewer 1:

Party ID*Linear Quarters to Election

“I do not think figure 6 helps the authors establish the validity of the partisan heuristics mechanism. Instead, the authors should just report the estimates from the regression tables (e.g., model A9 in table 1) and omit figure 6 altogether. On a separate note it is unclear why the authors interact party ID with the squared of “quarters to election”? Why is this preferred to party ID interacted with the linear (nonsquared) “quarters to election”? Are the results robust to this alternate interaction term?”

- The squared interaction term is preferred because both of the theories that include elections posit that it has a non-linear relationship with forecast errors (the absolute error should increase only close to elections). This is discussed on page 18. There is substantively no difference in the results for interactions between president party ID and the linear or polynomial version of the quarters to election variable. In the Supplementary Material we have added a note on this and include supporting regression output.

- Following the reviewer's recommendation we have removed the left panel of Figure 6.

FOMC Models:

“Do FOMC models stay the same over time? With the exception of a change to the FRB Global Forecasting model, do the econometric models underlying the Greenbook forecasts stay the same over time?”

- We have added at the end of the Supplementary Material a brief history of Fed Staff models over our observation period. The main finding in this history is that the models are altered over time in significantly smaller ways than the shift to the FRB Global forecasting model. In our research, we did not find a specific timeline of these tweaks, so it would be very difficult to develop a variable or variables that directly capture each small shift. As such we focus on the major shift to the FRB Global model.

“On a related note, do the authors have any sense how often the consensus forecasts depart from the predictions from the econometric forecast models?”

- There is a major empirical problem with separating out how much of the forecast error is model error and judgmental error. Only the consensus forecasts (model forecasts + judgmental adjustments) are made publicly available. So unfortunately it is currently impossible to directly observe model vs. judgmental sources of error. We have added a footnote (Footnote 4) describing this issue.

Mid-term Elections Robustness Check:

- In the Supplementary Materials we also examined a quarters to election variable that included mid-terms as well as presidential elections. This variable was, like the presidential election variables, statistically insignificant. Its interaction with presidential party ID was also insignificant.

Shocks

“Forecast errors frequently result from a variety of exogenous shocks which are completely omitted from the econometric models … It would be useful, as a robustness check, to control for oil prices, changes in productivity, and major external events (e.g., number of foreign wars in a given year) as additional controls.”

- We agree that the inclusion of potential shocks to inflation could increase the certainty of our results. We therefore examined a number of different shock indicators including absolute inflation level (both in the quarter being forecasted for as well as the quarter before the forecast was made), West Texas crude oil price changes, labor productivity changes, and the total number of armed conflicts per year. We looked at models with both the standardised and absolute inflation forecasting error variable as the dependent variable. The presidential partisan ID finding did not change when we entered the shock variables. A write up of these

variables and findings is in the Supplementary Materials.

Reviewer 2:

“Clarify what the length of forecast means (p.10)”

- We removed the term 'forecast length'. Instead we now use 'forecast age' throughout and define this more clearly in the text (i.e. how many quarters before the forecasted quarter a prediction was made) and formally. The definition is found on page 15.

“Introduction should tell readers more about why we should care about inflation forecasts. Inflation forecast bias is not necessarily the bread and butter of every political scientist, so we should understand (i) why it is important; (ii) the literatures that are relevant to the question and (iii) how has it been studied before.

“Give a sense of what other information the FOMC uses in making decisions. How important is the Greenbook data/information?”

- In response to these helpful points, throughout the paper, and especially in the first section (p. 3-4), we have tried to strengthen the case for why studying bureaucrats’ inflation forecasting errors is important. Adolph (2013) argues that Greenbook forecasts have a large effect on interest rate decisions. This is likely because there is a considerable delay from the time that interest rates are set to when inflation changes, so expectations are a very important part of monetary policymaking (Goodhardt 2001). A reflection of this is that a considerable amount of time is spent in FOMC meetings discussing expectations, rather than desired policy outcomes (Romer and Romer 2008). Greenbook forecasts are given to the FOMC before these meetings in written form and are presented by the staff at the beginning. They directly frame the members’ discussions and therefore expectations and ultimately policy choices. Inaccuracies in Greenbook forecasts could influence FOMC members to

make non-preferred interest rate decisions. This is why having highly accurate Greenbook forecasts and why studying inaccuracies is so important.

"Figure 5 – it seems that expected errors for Republicans get closer and closer to zero as the lag of the forecast is closer to present time. Why may that be the case?"

- First it is important to note that though the mean does get closer to 0 the closer the forecasts are in time to the forecasted quarter, there is a fair amount of overlap in the simulation distributions. So we should be cautious about over interpreting this result. Second, to the extent that forecast errors do get closer to 0 the closer in time that they are made, this is what we would expect. It is easier to make accurate forecasts the closer in time predictions are made to the things they are predicting. Current economic conditions are more closely related to conditions in the near past. It's notable that, though closer to 0, forecasting error is still estimated to be below zero during Republican presidencies, i.e. under-estimated as the heuristic theory predicts. We would be happy to add a discussion of this to the paper if the Reviewer believes it is necessary.

"If the heuristic explanation is right – would there not be an expectation that there is also a relationship between president party ID and unemployment forecast errors? Why is the lack of evidence on unemployment bias, supporting evidence for inflation forecast bias (p. 23)?"

- The last sentence of this section was poorly worded and we appreciate this comment alerting us to this. The idea of running the models with an orthogonal dependent variable as a robustness check was to examine if some unobserved time period specific factor--other than partisan ID--is creating forecast errors across a range of quantities and is not specific to inflation. We did not find this to be the case. A fuller discussion of this issue has been added to the Supplementary Materials.

- In particular, we added a discussion positing why presidential partisanship is not associated with unemployment forecasting errors. In

sum: (a) a heuristic only causes systematic errors when it does not correlate with the process behind the quantity being forecasted (e.g. Democratic and Republican presidents' effect on unemployment may actually be different in the direction and magnitude anticipated by the heuristic). Alternatively (or additionally), (b) dropping a handful of outliers the range of the magnitude of standardised forecasting errors is much larger for inflation than unemployment (see Figure 7, right panel). This suggests that forecasting inflation is much more difficult than forecasting unemployment. Possibly this is because employment is stickier than prices. As the heuristics literature suggests, in this situation presidential partisan heuristics would be relied on more when forecasting inflation than unemployment. The result of a heavier reliance on heuristics could be systematic inflation forecasting errors. More work, outside the scope of this paper is needed to disentangle these possibilities.

"The treatment of the effect of partisan politics on inflation and output seems quite thin. Examples are on p. 4 (including footnote 3)."

- We have greatly deepened the discussion of the partisan politics on inflation. See above.

"[A]uthors introduce jargon that is under explained, e.g. "rationality of forecasts" (p.3); behavioral economics literature (p. 24)"

- We have removed the jargon 'forecast rationality' and 'behavioural economics'. We now use more common sense descriptions of these concepts.

"Some statements are baffling, like the rationale for choosing the dependent variable (p. 9). Why is the assumption that central bankers are inflation averse a justification for the measure? This paper seems about making understanding in more detail the decision making process within the Fed, so we need to know about which "central bankers" is this story about:

the Fed staff, the FOMC?"

"We need to understand who is inflation adverse at the Fed: the staff or the FOMC? How do staff relate to the FOMC?"

- We have clarified the text (p. 13) such that the reason inflation forecast accuracy is the most relevant to study is not because central bankers are inflation averse, but that they are widely believed to be primarily concerned with price stability (e.g. Cukierman, Webb and Neyapti, 1992; Mukherjee and Singer, 2008; Tillmann, 2008). If this is true then inflation forecasts would be the most substantively important forecast influencing policy decisions and systematic errors would reduce their ability to achieve this goal.

- See above for a discussion of how the staff relate to the FOMC.

"To what extent is it reasonable to apply a framework of "relatively poor information societal actors" to Fed staff (p. 4 and again in conclusion). This would need some justification."

- We have removed this sentence from the paper as it was an error and, more importantly, expanded our discussion of the heuristics and monetary policy literatures on how *experts* do in fact rely on heuristics for forecasting in situations where there is high uncertainty about the quantities being forecasted.

"Government expenditure, debt and deficit are included because these may be some of the policies on which Democrats are expected to differ from Republicans. What would other policies be? Crucially, how about inflation expectations? We know that inflation expectations become unhinged in high inflation environments, which makes it hard to control and possibly predict inflation. It seems to me that authors should use past inflation levels and inflation variability in their models."

- Please see the final note to Reviewer 1 regarding our inclusion of the

absolute inflation measure as a control. This issue has been addressed in the revised version's Supplementary Material.