Referee Report on Manuscript Number 1232

Title: "Fairness and the Disinflation Puzzle

## **Brief Summary and general assessment:**

The author analyzes a credible disinflation in a shirking model, where wages are considered fair as long as they match past inflation. Workers' opportunities outside the firm are measured by the job finding rate rather than by the unemployment rate. This latter feature makes the model distinct from the approach by Driscroll and Holden (2004). The author shows that this approach leads to large costs of disinflation, thereby solving the disinflation puzzle.

In my view the paper is sloppily written. Neither the research question, nor the methodology are properly explained. The author does not make clear, why his or her approach is important and how his or her approach helps reconcile the model with the data. Furthermore, a thorough proof-read is necessary. There are many formal (and language) mistakes such as wrongly-spelled author names, random typos, and inconsistencies in the math notation (more on this below).

In general, the paper touches on an interesting topic and - in line with many authors, especially in the behavioral economics literature - highlights insights from behavioral economics, which help explain puzzles that mainstream theories cannot explain. However, up to this point, the paper lacks a clear recurrent theme as well as proper motivation for its assumptions and explanatory power. The model and the analysis have to be put much stronger into context wit a broader range of the literature on this topic.

## **Further comments:**

1. Without having read Driscroll and Holden (2004), the abstract does not yield any useful information about the research question or the methodology.

2. The previous comment touches upon a general and rather severe problem of the paper. Namely that without detailed knowledge of Shapiro and Stiglitz (1984), Kimball (1994), and Driscoll and Holden (2004), it is impossible for the reader to follow the author through his or her paper. This holds true for the general methodology and intuition as well as for specific situations. To give an example for the latter, in the abstract as well as in the introduction the author refers to a "much larger effect" than Driscroll and Holden (2004), without specifying what kind of effect it is (or even what general research question or topic it is) that Driscroll and Holden (2004) analyze. Another crucial example can be found on page 11, where the author states that "the expected wage... is obtained in a way analogous to Kimball (1994)." The author only presents the resulting wage, but not how it comes about. This is of high importance, though, as the wage determination is crucial for the results. However, the author does not elaborate whether the wage determination is a bargaining problem, a simple combination of previous equations, or something completely different. Without going into the paper by Kimball (1994) the reader has no chance to understand how the respective wage is determined. A description of which equation are put together in what way as well as what is done to these equation is indispensable in this situation. These are only two of many more situations. A careful proof-read with respect to the text's ability to transfer all the crucial knowledge to enable the reader to follow the line of reasoning without going into different papers is necessary.

3. The introductory section as well as the literature review need to be fully rewritten. The introductory section does not set up a clear research question to begin with. In order to understand the research question, the reader once again has to go back to the three cited papers above and imply from their contents as well as the comments in the text, what the specific research question really is. Only then, it makes sense to lay out in detail the methodological contribution of this paper. The particular approach chosen (fair wages plus the definition of the reference wage by Driscroll and Holden (2004)) needs to motivated much better. It would be useful to see some empirical evidence that the reference wage determination process is supported by empirical evidence. The author simply states that he uses the approach from Simonsen (1988) and puts it in Driscroll and Holden (2004). Why it should make sense to combine these two and what the expected result (or the contribution of this approach) is remains uncommented up to this point. Therefore, this part needs a proper motivation for the methodology and the assumptions chosen. Especially, the assumption why fairness matters needs to be thoroughly discussed.

4. Similar problems arise with the literature review. To my understanding, Section 1 ("The Disinflation Puzzle") is redundant in its current form as it does not add any significant information to the discussion of the paper. This section does not guide the reader through the literature, as it picks some random studies and leaves out others. Discussing indexation and the New Keynesian Phillips curve could be done on much simpler grounds than provided in the paper. Whether or not indexation can be observed on a quarterly frequency is an uninteresting question in the sense that indexation in any form is not an empirical phenomenon (see, e.g. Babecký et al. (2010) among many others). Indexation is commonly motivated alone by its merits for improving the empirical fit of dynamic stochastic general equilibrium models. Along similar lines, I find the discussion about sticky information little persuasive. Here, there is ample empirical evidence which could be cited, as e.g. Coibion (2010). The evidence cited in the paper is in no way persuasive, especially, since it neither proofs nor disproofs the existence and importance of sticky information. Thus, a simple sentence about there being mixed evidence on this issue and some sources would be more convincing and actually sufficient. Furthermore, there are some important sources missing. Costly disinflations in New Keynesian models arise with all approaches that lead to backward-bending longrun New Keynesian Phillips curves. Approaches that show this phenomenon can be found by Ahrens and Snower (2014), Graham and Snower (2008, 2011), Snower and Vaona (2008), Vaona (2013a, 2013b) and many others. A brief discussion of this issue would be useful.

5. The author seems to use the expressions fairness, reciprocity, fair treatment, and norms interchangeably. If this is so, the author should state that. Otherwise, it would be useful to define, what is meant by these concepts, as there are many different interpretations of all of these terms in the literature.

6. Section 2 should start with a brief overview of the model. The utility functions (1) and (2) simply fall from the sky, without being introduced or motivated. The remark concerning consumption habits is redundant. Furthermore, an explanation is needed, why it is worth noting that habit and peer effects are not considered. Are these consid-

ered in the original framework? Habit persistence is a feature of Danthine and Kurmann (2010), but in their utility function (in terms of consumption habits, while consumption is not even present in the model of this paper) not the effort function, which is the part that this paper takes from Danthine and Kurmann (2010). So I do not see the point in mentioning habits at all. In general, many equations and expressions seem to come out of nowhere and completely ad hoc. Has any other paper used these functions before? As I see it, they are a mix of Shapiro and Stiglitz (1984), Kimball (1994) and Danthine and Kurmann (2010). This should be stated. The author claims that the model and the notation is based on Shapiro and Stiglitz (1984) and Kimball (1994). However, the notation is much closer to (but not identical to) Driscroll and Holden (2004). This should be clarified.

7. The math notation is rather complicated and the analysis is not always straightforward. Especially, equation (5) is hard to read, as  $\eta$  seems to be misplaced. The parameter  $\theta$  in equation (9) on page 11 is not defined in the model section. For the reader the parameter remains a mystery until it is introduced in the calibration section on page 17. Similarly, the author presents an "alternative" interpretation for the parameter  $\phi$ , but supplies no initial interpretation for the parameter. Also, it is not clear what the underlying assumptions are that lead to the result of equation (6). Equations (7), (8), and (11) could be improved by using fractions. They are particularly complicated to read. This is especially true for equation (11). Further, it would be useful to have an argument, why the author can skip the superscripts identifying the firm and the worker (e.g., identical workers and identical firms). Finally, on page 8, why small letter "w"s, when everywhere else it is in capital letters? Are these log values? The same applies to page 20. According to the way the growth rate on page 20 is set up, they should be. But this should be stated.

8. It would be nice to learn, why the author uses past expectations for some, but not all current values. of course, this is supposed to capture period t - 1 judgements, but without further explanation it is hard to reconcile the functional forms with past judgements. In general, many equations could use some more intuitive explanations.

9. On page 12 the author refers to "price level higher than what would prevail in the previous paragraph." However, there is no price level in the previous paragraph. So, which price level does it refer to?

10. I do have a problem with the way the author summarizes his or her model. The author says that the steady state simply acts as a reference point against which shortrun deviations are measured. I agree with this interpretation. What I do not agree with is that the author derives all the steady state equations in detail. On the other hand, the "equations defining the dynamic short run equilibrium ... are derived in the appendices". In my opinion, the detailed presentation of the steady state within the text, which is not part of the analysis, is redundant or at least too lengthy. The really important part, the dynamics and their intuition, however, are put into the appendix. I do not see why this way of presenting the model makes sense. The author should carefully think about what is important and what not.

11. In the beginning of Section 4, the author describes what he or she is NOT doing in the following. This information is completely redundant or at least highly exclusive.

12. In section 4, the author refers to realistic costs of fairness, without elaborating on what this means. Is it the same as the price of reciprocity?

13. Concerning the calibration of the simulation exercise, the author claims to have summarized all parameters in Table 1 in the appendix. Unfortunately, Table 1 lacks the values for productivity growth g, which is given by 0.5 in the text. There is no information, whether this value is common or not and where the value stems from. Also, the author refers to visual proof given in Figure 1 of Shimer (2007), without presenting the figure in the paper. A footnote attaches to this citation even gives an explanation how to read the figure. However, it is up for the reader to download the paper by Shimer to be able to see the figure and to fully understand the information the author is giving. Concerning the percentage of monitored employees, the author chooses 90%, which is in the middle of the range of values chosen by Costain and Jansen (2009). However, this information does not tell the reader whether it is in the middle of 89% and 91% or 80% and 100%. Clarification would be helpful.

14. Where does the assumption comes from that in a credible disinflation the reference remains fixed for four quarters? How does that fit together with the assumption by Driscroll and Holden (2004) and how is that comparable to rational expectations?

15. The author assumes that wage contracts hold for three years. This is at odds with most empirical evidence on wage stickiness. According to Taylor (1999) and many others, wage contracts are written on average for approximately one year. How do the results change when going from three-year to one-year contracts?

16. Finally, there are a few typos, which should be corrected:

Page 4, paragraph 4: The authors are Smets and Wouters (2007), as also cited in the references.

Page 4, paragraph 1: The authors name is Danthine. This needs to be corrected also in the reference list. Also, the paper got published in 2010.

Page 17, section 4.1, line 4: There are two periods after Table 1.

Page 22, line 1: There is an  $dx_t^R$ . Is that a typo? Otherwise define this variable.