Author's response to reviews

**Title:** What determines income inequality between French male and female GPs - the role of medical practices.

**Authors:**

Magali Dumontet (mag.dumontet@gmail.com)
Marc Le Vaillant (maleva@vjf.cnrs.fr)
Carine Franc (franc@vjf.cnrs.fr)

**Version:** 3  **Date:** 18 July 2012

**Author's response to reviews:** see over
Revision note

MS: 1332299601694679

*What determines income inequality between French male and female GPs - the role of medical practices.*

Magali Dumontet, Marc Le Vaillant and Carine Franc
BMC Family Practice

Dear Editor,

Many thanks for allowing us to revise the paper and we are pleased to send you our revised version.

We considered all the comments proposed by the two reviewers and responded to all requests of clarification. In the following pages, we explain precisely how we have responded to each comment of the two reviewers (separately). But first, we tried to simplify considerably the manuscript as suggested by one reviewer and by you. In this line, we also were careful to your own suggestions (highlighted in the text): for instance, we simplified the abstract by avoiding reference; we rewrote all sections (or almost all) of the paper in order to simplify and to be understandable by a non-economist reader (we moved the technical part of the method in a dedicated “appendix”). We also removed the redundant parts particularly when presenting the results.

We hope that the revised version of the paper will be found satisfying.

We are obviously willing to answer to any additional questions or suggestions.

Thank you again for the comments that were helpful to improve the manuscript.

Sincerely yours,

Carine FRANC, PhD
Reviewer's report
Title: What determines income inequality between French male and female GPs - the role of medical practices.
Version: 2 Date: 29 May 2012
Reviewer: Anne-laure Samson

I - Major Compulsory Revisions

1. Comment on the background: I think the authors should motivate more their analysis: why is it important to study male/female GPs income differentials in France? This is not presented in the introduction. For example, maybe the authors could emphasize that there is currently a rapid feminization of the profession in France (and give figures). Female GPs working less than males, it may lead to potential problems for the regulation of ambulatory care in France and especially the regulation of the supply of physicians. It is therefore important to understand the reasons for this income differential.

   - As suggested by the reviewer 1, we present more precisely the context of the recent feminization that may crucially impact the provision of care. See Page 2 Lines 26-32:

      "In France, as in many other OCDE countries, there is a recent but strong feminisation of medicine and particularly of general practice []. In 2009, 29% of GPs were women vs. 13% in 1983 (Eco santé) and more than two over three of medical students who chose general practice were women. Numerous studies showed that women GPs usually work less than men (number of days per week and number of week per year) []. This may lead in the short term to potential problems in terms of both access to ambulatory care and general practice regulation. It is therefore essential to understand the determinants of GPs provision in terms of quantities and services provided and thus the determinants of GPs income gap."

2. The authors should detail more the different results obtained:
   a. The descriptive statistics (table 2) are not commented at all (p.5). The authors should give the main results of this table and the main differences between male and female GPs (and for example say something about a nice result that could motivate the empirical analysis, which is that male GPs work 48% more than female GPs but only earn 34% more). I think the 26% difference in income (obtained from the regression) is stated in the wrong paragraph ("male/female differences"), a paragraph that only deal with descriptive statistics.

      - The 26% of income difference is in the correct section since it comes from descriptive statistics [(Female income-Male income)/ Male income)] and not from regression. Our presentation was not clear enough. We carefully improve it such that (See page 5 Lines 124-137):

         "The descriptive statistics are presented in table 2. GI is on average 80 788 for female and 109 048 for male corresponding to a gap of 26% (taking the male income as reference). This income gap reflects partially the gender differences of workload: for example, women were found to perform, on average, 33% fewer services than men (consultations and visits). Moreover, women work significantly less in terms of days worked per week, of hours worked per working day and report to take more vacations than men. It is interesting to note that gender differences in workload appear to be higher than income gap.

Male and female GPs also appear to provide different contents of services: for example, 60% of female GPs report having frequently performed gynaecologic follow-ups vs. 24% of male GPs whereas male GPs report performing electrocardiograms, minor surgeries, and traumatology more often female GPs.

Being coherent with standard literature, descriptive statistics show significant differences between male and female GPs with respect to their own characteristics like age (p<0.01) and length of experience (p<0.05), as well as to their workload in terms of number of medical services (consultations and visits) (p<0.01),
number of days worked per week (p<0.01), percentage of visits (p<0.05) and finally to the percentage of children in their patient list (p<0.05).

b. Results of the estimates are not new and can already be found in the literature (p.5) but I think the authors should explain them a bit more: i) the experience effect shows that income increases until 24 years of experience and then decreases this could be an age or a generation effect; ii) how to explain the “rural” effect? The “sector 2 effect”? I think the paper should be presented slightly differently because the expected effects that the variables included in the regression should have, are very well described (p.11-12) but the estimated/obtained effects are not. Maybe the authors should combine both parts of the paper (specifications p.11-12 and estimates p.5) and give more importance to the interpretation of the effects obtained in the paper.

- As suggested by reviewer 1, we gathered estimates and results sections. We also added some comments on estimates to clarify the interpretations of our results. See Page 6 lines 140-163.

"The ordinary least-squares estimates of GI are reported in table3: the first column indicates the estimates for all GPs’ income regardless of gender; the next two columns present the respective results of estimates for men and women.

A Chow test [] indicates that the null hypothesis of equality of all coefficients in the male and female incomes models can be rejected at 5 percent level, meaning that income structures for male and female GPs are different. Considering all GPs, we show that length of experience (number of years in private practice) has, on GI, a significant positive effect (p<0.01) that decreases over time (p<0.01). GI increases during the first years of career then stabilises and decreases on average from the 24th year of experience. This well-known result could be due to an age effect and/or a generation effect. This is in accordance with previous studies []. Moreover, we find that GPs practicing in rural regions have a higher income than those practicing in urban areas (p<0.01) [], this could be explained by different patients profiles and/or by the type of services provided. The fee for service payment theoretically implies a direct relationship between private practice income and the number of services performed. As expected, the number of services performed by GPs has a significant positive effect on GI (p<0.01): an increase in the number of services by 100 results in an increase in GI by 2%. Salaried activity appears to have a positive effect on income (p<0.01). GPs group practices has also a positive effect on GI but this effect has low significance (p<0.10). Finally, being in ‘sector 2’ corresponding to GPs allowed to charge extra fees, has no significant effect on GI, note that we control by the number of medical services provided by the GP [].”

Comparing the estimates of the explanatory variables effects on income for male and female GPs, we get some differences: for instance, considering the length of experiment, an additional year in private practice has a positive effect for all GPs but this effect is higher on income for women than for men. We also observe that when all other factors are equal, women have a higher GI when they offer an additional consultation or visit: this could occur because of the content of the service. In the same line, the impact of having a salaried activity in addition to private practice is positive for all but higher for female GPs than for male.

c. Results of the decomposition: interpretation of the constant (p.7 and p.9): could the constant also reflect discrimination towards female GPs, for example from some kinds of patients (in rural areas, or older patients) that are less likely to see a female GP?

- To take into account this interesting remark, we added a sentence in the “Results” section to explain that a part of the constant effect gap may also reflect partially a discrimination effect (See page 8, lines 186-189 and page 10, lines 234-239):

“Finally, the constant differential is the third term in the decomposition and is equal to 0.731. This term is difficult to interpret as it gathers effects of unobserved differences in returns and differences in the means of
unobserved variables and also of differences in the returns on the omitted variables like discrimination effects [\[...\]].

“The differential between the estimates of constants in the two models finally explains a large part of the incomes gap. This differential is derived from unobservable characteristics (preferences, risk aversion, etc.) and then, could be the result of either the differences in unobservable characteristics -for instance, Rizzo (2007) showed that male and female physicians appeared to reveal different ‘target incomes’ [\]$-$ or differences in the estimated coefficients of these unobservable characteristics like discrimination effects (for instance, some kinds of patients in rural areas, or older patients that are less likely to see a female GP).”

d. Results of the decomposition: could the authors interpret a bit more the portion attributable to characteristics? For example I am quite interested in the result that female GPs have an advantage compared to all GPs in terms of belonging to the sector 2 and in terms of location. What does it mean? What are the consequences in terms of public policy? (when female GPs are sector 2 GPs, they overbill more than male GPs for example?) Concerning this second part of the decomposition, I think the authors should make clear that it is an unexplained part of the income gap, and therefore I think only hypotheses can be made.

- Even if these results appear to be very interesting to discuss, the corresponding estimates are not significant. This is probably due to the very low number of women practicing in sector 2 and located in rural areas. We decided to keep these variables in the decomposition because they were significant in the regression meaning that they contribute to determine GPs’ income, but there is no difference in their respective effect on income for women and male. As suggested by reviewer 1, we also clarified which part of the decomposition is unexplained (See page 7, lines 166-170):

“This total gap corresponds to the sum of three parts: one part explained by the average differences in characteristics called explained part, a second part, called unexplained, corresponding to the sum of advantages and disadvantages of male and female GPs (comparing to all GPs population) for each of the explanatory variables and finally, the third part explained by the differential between the constants estimated for the two models.”

3. The authors should test the robustness of their results to the definition of the income variable: Could the authors check if their results are maintained when they use an alternative definition of the income variable, i.e. an income variable that only contains incomes earned from the self-employed practice (exclude income from salaried activities, etc.). In fact, variables used in the regression and that explain the differences between male and female GPs are mainly variables that characterize the self-employed activity. As a consequence, it is not clear to me why the global income, including salaried activity, should be used.

- We checked the robustness of our results on three income variables, we added this comments in the methods section (Page 5, lines 118-121).

“We checked the robustness of our results on three different income variables (private practice turnover, GPPMP, GI) and because the results were very similar, we chose to only present the results obtained using the gross income (GI) variable (others results are available upon request).”

4. On the regressions:

a. Could the authors control for other demand variables? i) the % of CMU patients; ii) the level of medical density in each department (specialists and GPs) could reflect the potential level of demand. There is a strong heterogeneity in terms of medical density in the Midi-Pyrenees region that may not be entirely captured by the rural/urban dummy.
We tried to include demand variables (%CMU, % of children in their patient list, % of elderly in their patient list) and GPs density but probably because of the sample size they were not significant.

b. Because of its endogeneity, I understand why the authors do not include the number of hours worked in the regression (even if they could instrument it using the number of children under the age of 15 for example). But I do not understand why the authors do not present another regression that would explain the hourly income. In the section “background” they state that the literature finds significant differences and smaller male/female gaps using the hourly income instead of the annual income. In their conclusion, the authors interpret their results in terms of difference between male and female productivity. Another analysis that deals with hourly income would therefore be more convincing to say things about. (And I don’t understand well the argument given by the authors p.9 for not doing this kind of analysis).

We did not include hourly income analysis between male and female GPs because in France, physicians are not paid hourly and because of the decreasing of working time productivity, it is not appropriated to study the hourly income gap [15]. Moreover, by controlling for the number of medical services provided, our gross income can be interpreted as a productivity of the GP. We rephrased this point (Page 9 Lines 225-233).

“It is important to note that by controlling for the number of medical services provided, the GP’s gross income can be interpreted as the productivity of the GP (income for a given production). For instance, all things being equal, providing an additional medical service for a female GP generates a higher additional income, suggesting a higher productivity. Another important dimension of doctors’ productivity is the length of medical services provided: some studies agreed that female practitioners have longer consultations/visits than male ones [1]. By considering the length of consultation/visit as a proxy of quality [1], a higher marginal return for women may be viewed as a compensation for quality. Following this line, we could have studied the hourly income gap between male and female GPs but in France, physicians are not paid hourly and finally studying the hourly income gap was not appropriated [1].”

c. I understand that the authors are constrained by the small size of their sample and that they cannot include lots of explanatory variables at the same time. However, I am a bit embarrassed by the fact that the authors interpret their results by saying that women use more lucrative services (specialized procedures) but without testing it empirically (they mostly rely on descriptive statistics). However, the mix of services used by GPs (ie. the fact that they performing pediatric / gynecological follow-ups...) is not a variable included in the regression. Why? Why don’t the authors perform an alternative regression (if needed, without the variable “number of services” and maybe without the rural/urban variable because it may be highly correlated) but with the variables “gynecological follow-ups, pediatric follow-ups” (or a mix of these variables) in order to quantify the effect of these variables on the income gap.

We tried to include variables reflecting the mix of services in our regression -with the number of services- but they were not significant probably because of the sample size. Then as suggested by reviewer 1, we performed an alternative regression (without the number of services) adding variables on “pediatric and gynecological follow-ups”. Frequently performing Gynecological follow-ups has no significant effect on income as well as for male and female GPs. Nevertheless frequently performing pediatric follow-ups has significant effect on income for male GPs (p<0.05) and for female GPs (p<0.10) but this effect is stronger for women than for men. In order to simplify the paper we did not include these results but they are available upon request.
5. Income decomposition (table 4): Would it be difficult to give standard errors to all variables included in the decomposition to see which variables statistically influence the decomposition?

- As suggested by reviewer 1, we added p-value in table 4. (See page 17).

6. On the representativeness of the sample: The authors use a small sample of GPs, that contains information on GPs practicing in only one region. For the interpretation/potential generalization of the results, maybe the authors should give information on the representativeness in terms of medical density (p.10). Midi-Pyrenees is a region where medical density is quite high compared to the average level: what consequences does this characteristic have on the results and on their generalization?

Midi-Pyrenees is one of the largest regions in France constituted of highly urbanized areas (around Toulouse), rural and mountainous areas that are characterized different level of medical density (and not only in terms of GPs). We do not observe the “local medical density” and we implicitly assume that the variable “area” also captures the differences in medical densities.

II – Minor Essential Revisions

7. Minor comments:

- p. 5: Oaxaca-Ransom? Corrected
  - As suggested by reviewer 1, we replaced our reference by this more recent reference (Page 8, Lines 192-194).

“This result is comparable to the wage gap measured for salaried workers in France, which was 25.3% in favour of men [1].”

- p. 8: The authors refer to Delattre and Dormont (p.8), but I don’t think they should use the term “target income” because the authors do not test for this hypothesis and do not have data to test for this (and the authors are proposing a utility maximization model, rather than a model consistent with the target-income hypothesis). More generally, I think the authors should be very cautious when they use the term “target income” to interpret their results (p. 8) because they do not test whether female GPs have a target that they try to reach (note that, in the literature, it has also been shown that female GPs are less sensitive than male GPs to their target income, as shown by Rizzo and Zeckhauser).
  - As suggested by reviewer 1, we are more cautious with the term” target income” and we rephrased this point (see page 9 lines 217-224)

“Another additional explanation may be derived from the induced demand theory: Delattre and Dormont (2008) showed that facing an increase in medical density and thus being subject to rationing, doctors tend to compensate by increasing the volume of treatments performed per consultation or visit [1]. Our results may also be interpreted by following a similar intuition: female GPs who appear to choose to work less or undergo a professional time constraint, compensate for this rationing by increasing the volume of treatments performed per consultation or visit [1]. Nevertheless, if it is rather well-known that female doctors generally see more women and children [Error! Reference source not found.] this can also result from patients’ demand and preferences.”
- p.11: the creation of the income variable is not very clear to me: why do the authors need to use an expense rate from the ministry of health while they have most expenses in their database (rent for the office, …)? I also don’t understand why the expense rate is so low (25.3% while I thought it was about 45%). Maybe the authors should explain a little more the construction of the variable.

- We used an expense rate from the ministry of health because in the database there were many missing response for the GPs expenses and to clarify the creation of the income, we rephrase this point (Page 4 lines 94-98).

“The private medical practice turnover and fees for a locum were collected in the study. The doctors used data from official documents (Système National Inter-Régimes -SNIR- data for the turnover and Declaration number 2035A / tax return). The average operating expense rate has been estimated by the French ministry of Health to 25.3%. This rate excludes taxes and social security contributions and corresponds on average to personnel expenses, acquisition, rent, transportation fees [1]. ”

- p.13: the description of the reference is unclear: what is the reference point in the Oaxaca-Blinder? What is the usual reference in the Oaxaca-Ransom decomposition? How sensitive is the decomposition to the choice of this reference?

- As suggested by reviewer 2, we decided to simplify the paper and the presentation of the method and we removed this point.

- p.21: use the term experience and not age in table 4.

- We corrected this mistake.
Reviewer's report
Title: What determines income inequality between French male and female GPs - the role of medical practices.
Version: 2 Date: 1 June 2012
Reviewer: Madelon Kroneman
Reviewer's report:
Health Policy Review

Major compulsory revisions
1. Since the scope of BMC Family Practice is not on gender differences but on aspects of primary health care, including clinical management of patients, professional training, shared decision making, and the organisation and evaluation of health care in the community, I would expect that the discussion should focus on the meaning of income differences for healthcare provision and expenditure.

2. Also, the discussion should be phrased in such a way that it is comprehensible for non-economists as well, since a large part of the readers will have no economic background. As a non-economist, I had a hard time understanding the paper.

As suggested by reviewer 2 (to answer to the first comments), we rewrote the discussion (and more broadly the whole paper) simplifying significantly the presentation and interpretation of our results. We particularly focused on the interpretations of our results in terms of the meaning of the gender income differences for healthcare supply by providing different assumptions about what may hide a higher productivity (higher return of an additional visit) of women doctors (see pages 8-10).

3. Furthermore, I think it is necessary to describe the methods and findings in a more concise way.

We rewrote the sections “Methods” and “Results” in a more concise way (see pages 3-8). We also decided to present the technical part of the method in an appendix. The idea is to give some clues on the properties of the decomposition for the readers who want to go forward into the model.

4. The authors have chosen to place the methods section at the end of the paper. Since this section contains important information to understand the results section, I would advice to place this section in between the Backgrounds and the Results.

As suggested by reviewer 2, we placed the methods section between the background and the results.

5. The Background section discusses in detail the previous research. I would suggest shortening this section and focusing on the common findings of these studies.

We rewrote the background section and we focused on the common findings in the literature (see pages 2-3).

6. The methods section and results section will benefit from some more clarification and clearer formulations.

See answer 3 (see pages 3-8).

Minor essential revisions
The comments are rather detailed, but some are essential for understanding the paper correctly.
Methods section:

7. Please avoid footnotes. The footnotes in this text sometimes offer important additional information and none of them are longer than half a line, thus including them in the text will improve legibility.
   - We introduced footnotes in the text.

8. Retroceded fees are fees for a locum? (p.11)
   - We replaced “retroceded fees” by fees for a locum (See page 4, line 93).
   
   “GPPMP= (private medical practice turnover – fees for a locum) x (1 - average operating expense rate)”

9. What do you mean with “The salaried activities and various indemnities are exclusively declarative data”.
   - We mention here that the doctors were asked to report data on the basis of official documents (tax returns document), but ultimately, this is leading to declarative data standard limits (standard reporting bias). Nevertheless, the sample appears to be representative of the French GPs.

10. “To specify incomes.” Do you mean here to explain or to estimate incomes? p. 12
    - Actually the meaning was “to explain income by estimates” (by estimated coefficients for all explanatory variables). The section has been simplified.

11. The authors use both number of services and working time as independent variables. Aren’t these highly correlated, implying one cannot include both entities in the analysis?
    - We only described the two variables in the descriptive statistics and of course, we did not include the working time -like the number of hours worked- in the regression because this will cause an endogeneity bias.

12. “We also expect that income decrease with having a specialised practice or a particular manner of practicing that represents 30% of the doctor’s practice” I do not understand what the authors state here? What kind of specialism is meant here? Can you provide an example?
    - Since the physicians under study here are GPs, who are generalists in nature, it is not clear to me what kind of specialism is meant here.
    - Here, we considered that GPs have a “specialized practice” if they perform activities such as acupuncture, homeopathy, dietetic, etc for more than 30% of their working time. Even if these activities are not considered as specialties, they may modify the content and the organization of the routine private practice. That is why we include this characteristic in the descriptive statistics analysis.

13. I would suggest to rephrase the sentence “patient profiles associated with radically different resources”. I think you mean that some categories of patients visit the doctor more often and with more complex problems.
    - As suggested by reviewer 1, we reorganized the presentation of our working assumptions and the results section, so this sentence has been also modified.

14. Please provide a legend with the formulas. (Explaining M and F for instance)
    - See appendix 2 page 19
15. The methods section does not address the statistical tests used for testing the significance of the differences.
   - As suggested, we made some clarifications and we reorganized the presentation of the assumptions and the results; in the table 3 we removed the test of difference in coefficient and we added the p-value in the decomposition. Because of these clarifications we did not report in the methods section the statistical tests used for testing significance of the difference in coefficient (See pages 16-17).

16. The Appendix with the variables could be elaborate:
   - Please provide a definition of what is considered a rural area
   - Please provide the way GPs are asked about their health. Is this done in a validated way?
   - Salaried(%) : % of total income?
   - Percentage of visits: what is the denominator?
   - What is frequently in the special services mentioned here? Is this a yes/no variable or is it continuous, just providing the number of these services? Similar question for the number of certain patients: what is “many”?  
   - Number of vacations: why use number instead of number of days or weeks?
   - What is considered a vacation? One day off, a week, longer?
   - Variables used in the decomposition of the difference in income are the same used in regression. In table 4, we added “p-value” in order to identify the significant results (See page 17).

17. Please give table titles that fully describe the table content. For instance: Table 4. Income decomposition should read: Decomposition of the difference in income between male and female GPs. Are only significant differences shown here? If not, why were not all variables included in the table. Please also use consequent terms for variables (location is the same as rural region?)
   - Variables used in the decomposition of the difference in income are the same used in regression. In table 4, we added “p-value” in order to identify the significant results (See page 17).

18. Table 4 is rather complicated and may benefit from clearer headings, for instance providing the same terms as in formula (1.4) and making clear that the third column is an addition of the former two.
   - As suggested by reviewer 2, we put formulas in the table 4 (See page 17).

19. Table 3: Please indicate why doctor’s practice variables are not included in the table p. 6
   - Now, in the section “Method”, we indicate why the variable “doctor’s practices” is not included in the regression (See page 5 lines 104-108).

“Firstly, we specified an econometric model (least square model) to identify the main determinants of income and that, among four groups of variables: GPs’ characteristics, workload, type of practice and type of patients (see appendix 1). Due to sample size, we decided to be parsimonious and not to include lots of explanatory variables at the same time. Using a backward method, we chose the most significant explanatory variables. At the end, explanatory variables can be grouped in three categories: GPs’ characteristics, workload and type of practice.”

20. What do you mean with “combine the experience effect”?
   - We sum the estimates of the experience effect (length of experience + length of experience squared) in order to simplify the reading of the results (overall experience effect) as we focus there on the gender difference of this effect (See page 17, table 4).

“Notes: We gather the experience effect (length of experience+ length of experience squared).”
Conclusion
21. I do not understand what the authors mean with the “adjective ‘mechanical’”
   • We rephrased this point (See page 10, lines 250-252).

“This even if workload is clearly an essential determinant of income and thus a determinant of the gender incomes gap, our results contradict our assumption as FFS does not reduce the gender incomes gap and there is a non perfect relationship between the provision of medical services and the income.”