Plain Assumptions and Unexplained Wizardry Called in Aid of *'The Fiscal Effects of Immigration to the UK'*

By Mervyn Stone (December 2013)
University College London, UK

Online Report: December 2013

55 Tufton Street
London SW1P 3QL

Civitas is a registered charity (no. 1085494) and a company limited by guarantee, registered in England and Wales (no. 04023541)

email: info@civitas.org.uk

Independence: Civitas: Institute for the Study of Civil Society is a registered educational charity (No.1085494) and a company limited by guarantee (No. 04023541). Civitas is financed from a variety of private sources to avoid over-reliance on any single or small group of donors. All the Institute's publications seek to further its objective of promoting the advancement of learning.

Plain Assumptions and Unexplained Wizardry Called in Aid of '*The Fiscal Effects of Immigration to the UK*'

Mervyn Stone

'Econometrics' is the application of a particular variety of statistical modelling to economic questions whose answer (if there is one) is not immediately obvious without such modelling. 'Econometricians' are scholarly academics who practice that specialization in the belief that econometrics can answer almost any question—when enough scholarly assumptions are admitted. Uninformed econometric wizardry can dominate the answers to policy-making questions in government, when they are posed by economists who may have forgotten (or are prepared to ignore) what they may have learnt in their statistics courses at university and when, in the machinery of government, mathematically complex questions are contracted out to econometricians. By the time the answers that government departments are then able to formulate get any media exposure, the assumptions that safeguard academic reputations receive little publicity—even if they are understood.

A nice example is the Nuffield Trust formula that NHS England is using to fund Clinical Commissioning Groups (www.civitas.org.uk/pdf/stoneseries). The formula's econometric assumptions have been understood and rejected by statisticians, but the rejection has excited little media interest. Nevertheless, my informal contacts with policy-makers at a recent NHS England workshop on CCG resource allocation suggest that there may be some weakening in reliance on econometrics.

A recent report, by Christian Dustmann and Tommaso Frattini [1] from UCL's Centre for Research and Analysis of Migration (CReAM), suggests that the wizardry may be weakening elsewhere. This report ('Cream' for short) is an ambitious and largely scholarly study, in which crucial assumptions (about how to share expenditures and revenue between immigrants and natives**)** are set out with commendable clarity—and are therefore open to a degree of critical comment by those who like to dig below the headlines and can handle the necessary source data. Econometric modelling was not invoked for the estimation of fiscal effects, although it may have been involved in the derivation elsewhere of the government statistics on which the estimation is based. The effects were simply estimated by straightforward accountancy of monetary figures from official sources, once the assumptions about shares needed to stitch the numbers together were fixed (humdrum calculations no more complicated than the simplest of household accounts do not require the rather forbidding algebra of Cream's equation (3).)

Where econometric hypotheses were still put to work was in the estimation of coefficients in some 'linear probability models' for the binary (1/0) data of interest to Cream—models rarely evoked by

[1] *The Fiscal Effects of Immigration to the UK*, by Christian Dustmann and Tommaso Frattini, Centre for Research and Analysis of Migration, 2013, http://cream-migration.org/publ_uploads/CDP_22_13.pdf (retrieved 29/11/2013)

statisticians but models that can be fitted quite simply by Ordinary Least Squares (OLS). The signs of the estimated coefficients in Cream's Table 3 are consistent with the estimates of fiscal effects and, in the Discussion, the authors are happy to present some of them as providing unqualified and impressive evidence of a "*differential in probability of claiming state benefits/tax credits or living in social housing* " between natives and immigrants. The differences are described as having 'impact' and the findings are warmly presented as supporting the fiscal effect estimates that openly rely on a number of subjective assumptions. Cream's Table 3 has obviously been written for a technically savvy readership not deterred by the significance (for them) of the phrase 'robust standard errors' and the asterisks associated with them. I will later explain the calculation of the results in Table 3 for readers with no more than high-school algebra, as a prerequisite for providing some missing but important qualifications and caveats. All of the calculations in Cream are heavily reliant on self-reported data from respondents to the quarterly Labour Force Surveys, which has non-response rate in excess of 10%. Cream's claim to precise estimation (p.3, quoted here in Conclusion) should have been tempered with knowledge of such realities.

Professor Dustmann may have wanted the assumptions in the calculation of fiscal effects, at least, to be clearly stated—if only to avoid the kind of unforgiving scorn that was heaped on the 2003 study he directed for the Home Office[2] as soon as its prediction became comparable with what transpired[3]. Questionable assumptions had delivered a confidence interval (5000 to 13000) for the prediction of the average net migration flow from 8 accession countries over the next 10 years—but the flow for just one year exceeded prediction by so much that a ministerial head was obliged to roll. Before the actual numbers came in, I questioned the assumptions in a Civitas piece[4] whose subtitle was *Technical exercise, honest study, or convenient obfuscation*? Dustmann presumably saw the assumptions as caveats without which no prediction would have been possible—which appeared to be his stance when the prediction was challenged by a House of Lords Economics committee.

A bad start!

Cream starts inauspiciously! The Abstract of a scholarly paper should be a faithful summary of what is in it—so that those with little or no time to read a technical paper (and that seems to be most of us) do not miss anything important.

For fiscal effects, Cream distinguishes five categories of residents—in households sampled by the quarterly Labour Force Survey (LFS) and interviewed in five successive quarters, between 1995 and 2011

---

[2] Christian Dustmann, Maria Casanova, Michael Fertig, Ian Preston and Christoph M Schmidt, *The Impact of EU Enlargement on Migration Flows,* Home Office, 2003, page 57, from http://www.irr.org.uk/pdf/rdsolr2503.pdf, (retrieved 29/11/2013)

[3] *Polish People in the UK - Half a million Polish Residents*, Office for National Statistics, 2011 http://www.ons.gov.uk/ons/rel/migration1/migration-statistics-quarterly-report/august-2011/polish-people-in-the-uk.html (retrieved 29/11/2013)

[4] Mervyn Stone, *Prediction of future migration flows to the UK and Germany*, Civitas 2003, from http://www.civitas.org.uk/pdf/EUmigration.pdf

1: Natives (born in the UK, excluding immigrants' native children)
2: EEAs (immigrants from the European Economic Area)
3: Non-EEAs (immigrants from elsewhere)
4: Recent EEAs (post-2000 immigration)
5: Recent non-EEAs (ditto)

without distinguishing

6: *Non-recent EEAs*
7: *Non-recent non-EEAs*

and necessarily excluding

8: *Illegal immigrants* (unlikely to be recorded in any Labour Force Survey).

So far, so good! Cream estimates the ratio of Treasury revenue to Treasury expenditure*,* for each year and for categories 1 to 5—but the Abstract mentions only findings that are favourable to immigrants (where the ratio is greater than unity). It does not report the significant feature that, for category 3, the ratio is appreciably less than unity in all 17 years. The omission is all the more surprising because the Introduction suggests that the paper was written to provide *substantive evidence on immigrants' fiscal contribution*—presumably as salutary correction of the Europe-wide poll showing that only 15% of people believe that immigrants receive less than they contribute (which may be because most Europeans see 'immigrants' as those who were born outside the EEA). Thus far, no reason is apparent why Cream should place any emphasis on the results for EEA immigrants, or even on post-2000 immigration.

Population changes

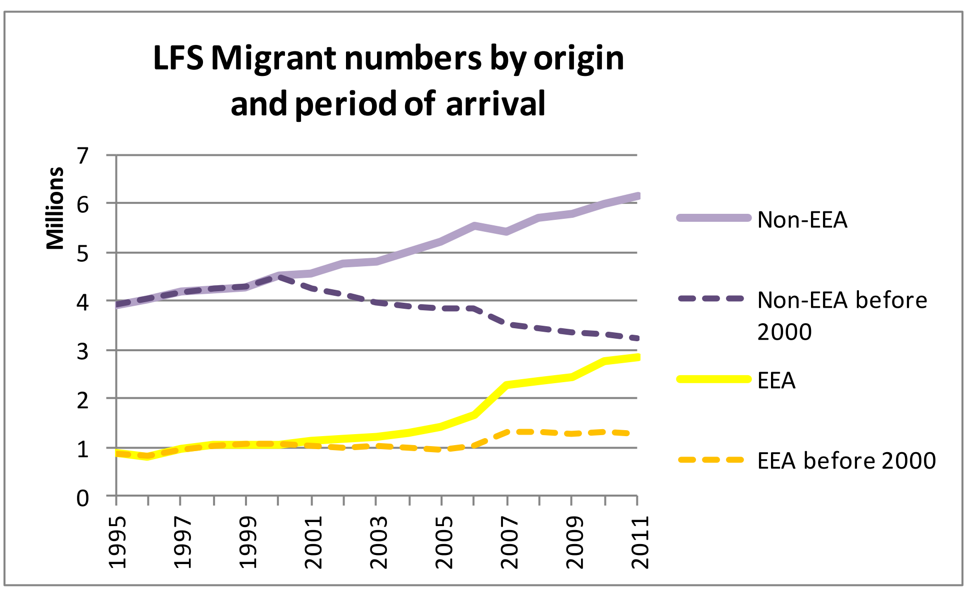

Figure 1

Figure 1 plots population estimates from LFS in Cream's Table 1a. To allow the correlated shifts in migrant numbers between 2006 and 2007 (from some LFS redefinition) to show up more clearly on an appropriate scale of single-digit millions, the Figure omits the remarkably constant 52 million for the native numbers—which exclude immigrants' native children. Native population is shown in Figure 2.

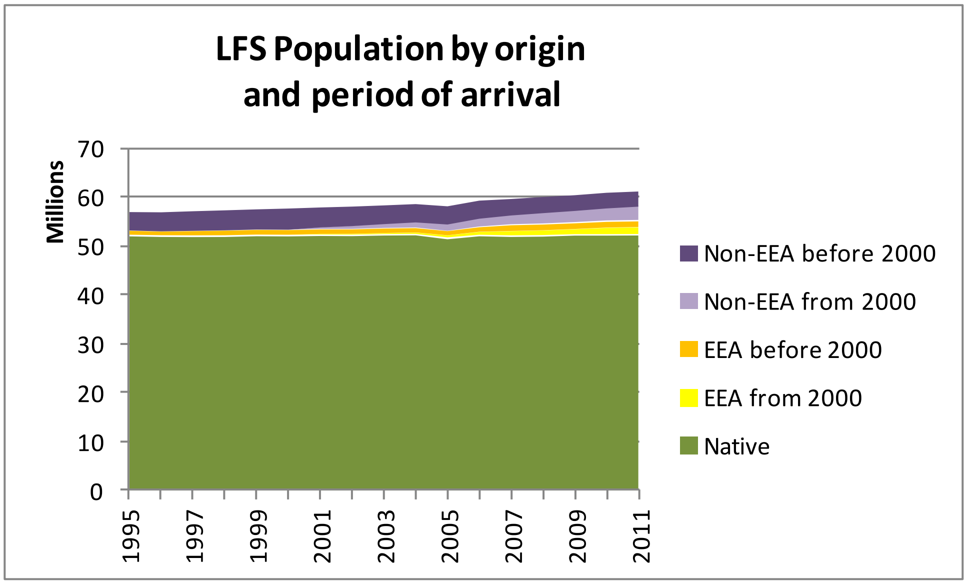

Figure 2

The Cream fiscal effect calculation

The Europe-wide prejudice about the net cost of immigrants does not, and could not, exclude any cost or benefit. So for any 'substantive evidence' to be convincingly corrective, it has to cover, for the UK, all the streams of financial exchange between government and the distinguished groups—however difficult that might be. Here are my edited transcriptions of Cream's Tables A1 and A2:

%	Expenditure	Basis of share between immigrant and native
23.3	'Pure' public goods	Number (No.) of over-16s *(average cost)*
7.5	'Congestible' public goods	No. of over-16s *(ditto)*
16.9	Health costs (except medical research)	No. and national *per capita* health cost, by age group
0.7	Compulsory education: pre-primary	No. of under-5 children
8.4	Compulsory education: other	No. of children over-5 and under-16
1.0	Further education	No. in further education
1.9	Higher education	No. in higher education
4.8	Sickness and disability	No. of sickness or disability benefit claimants *(actual recipients)*
13.2	Pensions	No. of pension claimants *(ditto)*
7.4	Family and children	No. of income-support or family-related benefits claimants *(ditto)*
1.0	Unemployment	No. of unemployment benefits recipients *(ditto)*

3.5	Housing benefits	No. of housing benefits recipients *(ditto)*
4.5	Personal social services	Nos. of sickness and disability benefits recipients, pension claimants, income support or family-related benefits recipients*(ditto)*: percentage share is the average of percentage shares of the three numbers
1.9	Law courts and prisons	No. in prison
1.1	Housing development	No. of social housing tenants
0.3	Immigration and citizenship police services	No. of immigrants
2.8	Other police services	No.
−0.2	EU transactions	No.
100.0		

%	Receipts	Basis of share between immigrant and native
44.7	Income tax and National Insurance	Total payments: actual tax and NI rates applied to LFS income
28.0	VAT and other indirect taxes	Total payments: ONS-group rates by household income decile
9.3	Company and capital taxes	No. of adults (but net ONS-defined receipts from foreign-owned shares)
4.2	Council tax	No. of households
4.3	Business rates	No. of self-employed
5.0	Gross operating surplus and rents	No. of adults (but 100% to natives, for the *marginal contribution* option)
0.6	Inheritance tax	No. of house-owners
−0.8	Income tax credits	No. of dependent children
4.7	Other	No. of adults
100.0		

Once the assumptions underlying the group shares have been accepted and the annual total expenditure and total receipts have been listed (which Cream does not do), the calculation of annual net fiscal contribution is straightforward (who needs the algebra of Section 2.2.2?). For categories 1, 2 and 3, Cream's Table 5 gives the totals of these contributions for the overlapping periods 1995-2011 and 2001-2011, and the totals for categories 4 and 5 for 2001-2011. It took not a little time to extract the messages in these totals. For instance, the deficits of £604b and £624b are (almost) telling the same story and the reader is faced with the totally unnecessary task of looking at what makes the difference between them, i.e. what happened before 2000.

Moreover, while all the totals are relevant to the management of the cumulative Treasury budget, it is arguable that, since the popular prejudice is clearly *ad hominem,* an informative rebuttal would include figures on a *per capita per annum* basis (i.e. per individual-year). The inclusion of population proportions in its Table 5 may be Cream's recognition of that distinction, but the authors leave it to the curious reader to do the calculation. Because Table 5 is (clearly) neither transparent nor inclusive of groups 6 and 7, we have here used its billions, together with the population figures of Table 1, to calculate directly comparable figures in an almost symmetrical table for all seven groups, which looks at the same figures from different angles:

Group	1995-2000 Total-net £m	1995-2000 Per individual-year £	2001-2011 Total-net £m	2001-2011 Per individual-year £
1: Native	19,591	63	−624,120	−1087
2: EEA	−203	−35	8978	436
3: Non-EEA	−17,356	−688	−86,820	−1471
4: Recent EEA	*	*	22,106	2732
5: Recent non-EEA	*	*	2,942	162
6: Non-recent EEA	*	*	−13,128	−1052
7: Non-recent non-EEA	*	*	−89,762	−2198

Note the necessary consistencies in column 4: 8978 = 22,106−13,128, −86,820=2942−89762 .

The popular prejudice against immigrants does not exclude category 8—illegal immigrants, whose numbers have been estimated at a million (who knows?). Cream did not try to estimate what effect they may have on different sorts of expenditure or on receipts from other categories of immigrant. A less specific omission is that, in Tables A1 and A2, Cream gives neither the share percentages nor the billions corresponding to the two 100% totals— which are needed to assess:

a) the relative importance in the national budget of the net fiscal contributions in Table 5 (or the derived figures in the table here);
b) their sensitivity to any particular share percentage.

Note that Cream presented the results of sensitivity tests in Tables 4a and 4b for just one expenditure (public goods), and that elsewhere Cream claims to have made only assumptions that favoured natives:

> ... when faced with an option about alternative ways of allocating fiscal costs to immigrants we have chosen throughout the paper to calculate a “worst case” scenario, from the immigrants’ standpoint, in the sense that the net fiscal impact of migrants is most likely to be more positive than our estimates suggest. **(p.4)**

The econometrics that Cream calls on to estimate putative ‘probability gaps’

Table 3 presents the results (pre and post 2000) of fitting five linear probability models—two models for whether or not a ‘randomly drawn individual’ (quoting Section 2.2.1) is recorded by LFS as a benefit recipient (of state benefits/tax credits), and three models for whether or not a household (randomly selected by LFS) is recorded as social housing. The models were probably fitted by OLS but I stand to be corrected about that. Since Cream’s Discussion gives prominence to the estimated coefficients in columns 1 and 2 of Panel B of Cream’s Table 3, I will focus on those coefficients (here re-tabulated for ease of reference):

Coefficients of the explanatory variable I (1 for immigrant, 0 for native) in models for the dependent variable D that is 1 or 0 for whether or not an individual is classified as a benefit recipient between 2001 and 2011

Natives are compared with:	Model 1: Fitting variables are Year-quarters and I	Model 2: Fitting variables are Year-quarters, I, Sex, Age and Age^2
4 + 5 Recent immigrants	−0.178 (0.001) ***	−0.084 (0.001) ***
4 Recent EEA	−0.201 (0.002) ***	−0.100 (0.002) ***
5 Recent non-EEA	−0.167 (0.002) ***	−0.077 (0.002) ***
Sample size N:	3,495,478	3,495,478

Robust standard errors in parenthesis. The sample 'proportion of native recipients' is 0.391 .
* Denotes significance at 10%, ** significance at 5% and *** significance at 1%.

This part of Cream's Table 3 calls for a number of observations:

(i) Cream does not say much about the provenance of the sample of size N other than to say that the units in the analyses were (presumably separate) '*individuals*' (p.10) and (with some ambiguity) that to '*increase the sample size, we pool the four quarterly waves in every fiscal year, which in the UK begins in April*' (p.16). I do not see how these statements correspond to the value 3.5m for N and the well-documented structure of the LFS sampling scheme. I think the units were not (cannot have been) individuals but probably LFS records—with about 4/5-ths replication of individuals (ONS says 80%). LFS tries to sample about 12,000 new households every quarter but gets responses from about 10,000, or about 440,000 different households for the 11 post-2000 years. (Each household is interviewed five times in a 'wave' of successive quarters and, for the ONS quarterly official statistics, five (not four) waves are pooled to make the quarter's records.)

If there had been 3.5m different individuals providing a record for Table 3 in these 11 years, that would be an average of eight per new household. If I were a gambler, I would be happy to bet against that—even without asking anybody in ONS about it. Counting the same individual five times is a reprehensible way of increasing the size of a sample—one that necessarily affects the shape of any statistical inference from Table 3 and also leads one to doubt the quality of the underlying analysis.

(ii) It is a pity that Cream's Table 3 did not include the potentially informative values of the traditional statistical measure R^2 of overall model performance (how well a model fits the data). See paragraph (ii) below.

(iii) The percentage 'significance' levels of the six coefficients are automatically determined by the ratios of coefficient to the 'standard error' used (the t-values in my Appendix). The rounded 'robust standard errors' are probably the result of a box being ticked in some software to request their 'robust' estimation.

(iv) Note that LFS sampling is of households not individuals, and that it is a hugely complex Office of National Statistics operation, with carefully-designed technical devices, adjustments and imputations that may be successful in eliminating biases, such as those from acknowledged variation in response rates of different sections of the population. Most users of LFS data do not feel that it is justifiable, for a single analysis, to do much more than quote estimates, without attempting to make any statement of their precision, especially when data is for several years. The estimates in Cream's Table 3 have the additional uncertainty of dependence on an empirical model that has not even been shown to fit observable but untabulated proportions.

(v) Perhaps for simplicity of presentation, Table 3 is not neutrally symmetrical with respect to immigrant/native status, and it is remarkable that the same figure 0.391 for the 'proportion of native recipients' (strictly speaking, the proportion of recipients among natives) appears to be used for both Models 1 and 2—to calculate the 45% of the Discussion as 0.178/0.391 and the 21% as 0.084/0.391. Unless there has been a remarkable coincidence, the denominators here would have been different if the 'proportion' 0.391 had been what it ought to be for this calculation—a putative 'probability' separately estimated in both models. There would be no point in any modelling, if its outputs were simply the observable proportions. That they are not, is clear from the explicit expression in equation (3) below, which shows that the estimate β^* of the 'probability gap' for Model 1 is not the difference of simple proportions for immigrants and natives. Are we therefore obliged to conclude that the authors deviated from neutrality between immigrants and natives in their calculation of 45% and 21% in the Discussion? (It is irrelevant that using a model 'probability' in place of 0.391 might have made little difference, when semantic hygiene is at issue.)

Model 1 can be exhibited as a slight simplification of Cream's equation (1):

$$D \approx \beta I + \tau_q \qquad (1)$$

by omitting the α in Cream's formulation (which does not change the model provided there is a τ_q in all 44 quarters). Cream tells us how it sees Model 1:

> When we regress our indicator variable [D] only on immigrant status [I] and time dummies [τ_q], the coefficient *indicates* a percentage points difference in the probability of receiving benefits or living in social housing between immigrants and natives observed *at the same moment in time*. This observation answers a question that is important for assessing immigration's fiscal cost: "Is a randomly drawn immigrant more or less likely to receive benefits (live in social housing) than a randomly drawn native, and if so, by how much?". (p.10) [My emphasis]

The right-hand side of (1) is regarded as a probability only because the probabilistic expectation of D (that it is trying to approximate) is the probability that D=1 (i.e. that an individual is a benefit recipient). It is because the least-squares fitting (explained in my Appendix) aims to get the right and the left as close as possible (overall) that it is natural to think of $\beta I + \tau_q$ as a probability—and β as the 'probability gap' between individuals with D=1 and those in the same quarter ('at the same moment in time') with D=0. If the least-squares estimation had been constrained (as it could have been) so that the estimate of $\beta I + \tau_q$ is in the interval 0 to 1 (which is where probabilities have to be), it would not have been OLS and the estimates would not have the unbiasedness property of OLS estimates. Did Cream check whether the estimate did fall in the 0 to1 interval for all quarters?

There is more to be established for Model 1, by a breakdown of the N = 3.5m 'individuals' (records?) used to fit both Model 1 and Model 2. The subset of n_q individuals (immigrants and natives in the q^{th} of the 44 quarters of the period 2001-2011) constitutes the 2 x 2 frequency table:

	I = 1	I = 0	Total
D = 1	a_q	b_q	$a_q + b_q$
D = 0	c_q	d_q	$c_q + d_q$
	$a_q + c_q$	$b_q + d_q$	n_q

The difference of proportions $a_q/(a_q + c_q) - b_q/(b_q + d_q)$ can be written in more informative notation as

$$\Delta_q = p_q(D=1|I=1) - p_q(D=1|I=0) \qquad (2) \ .$$

The last quotation of Cream makes sense only if the 'coefficient' does its 'indication' in every quarter and the random drawings of immigrant and native are done 'at the same moment in time' i.e. quarter by quarter—and if we think of $p_q(D=1|I=0)$ and $p_q(D=1|I=0)$ as 'binomial proportions' (determined by random drawing of a_q out of $a_q + c_q$ immigrants, and of b_q out of $b_q + d_q$ natives) corresponding to model probabilities that may vary from quarter to quarter but whose difference is the constant β in Model 1 that β* estimates. If a statistical analysis of the observable differences Δ_q revealed a variation from quarter to quarter in excess of what would be expected for drawing red balls with replacement from two urns with estimated proportions of red balls, how could that finding be reconciled with the sense of the quotation—and how would the parameter β even be definable? Any serious inconsistency would undermine the bold inferences that Cream draws from the results in Table 3. It is therefore regrettable that Cream does not tabulate the 44 values of Δ_q, which would be the starting point for a statistician with a penchant for Exploratory Data Analysis. Their statistical analysis would be able to reassure us that the constant parameter β in both models is not a figment of some econometrician's imagination—unless it really is!

My Appendix shows that, for the true-or-not formula of Model 1, the estimated coefficient β*, say, is the weighted combination of the differences Δ_q with weights w_q:

$$\beta^* = w_1 \Delta_1 + \ldots + w_{44} \Delta_{44} \qquad (3)$$

where the weights w_q (adding to 1) are proportional to $(a_q + c_q)(b_q + d_q)/n_q$ (which is the intuitively sensible product of n_q and the separate proportions of immigrants and natives).

Cream's Tables 1b and Table2a shows that there are wide differences in age profile, average age and average educational qualifications between immigrants and natives, especially for recent immigrants. To 'take account of' these differences (the customary parlance of most formula-makers), Cream has to use a more complex formula—and believe that it is a faithful representation of reality. This is how Cream made the case for the Model 2 extension of Model 1:

$$D \approx \beta I + \tau_q + \gamma_1 \mathrm{Sex} + \gamma_2 \mathrm{Age} + \gamma_3 \mathrm{Age}^2 \qquad (4)$$

where Sex is a dummy F=1/M=0 variable, and Age may be just what LFS records:

> Also of interest are comparisons between immigrant and native populations that are identical in some observed characteristics. For instance, one such question would be what the difference is in the probability of receiving benefits (living in social housing) between immigrants and natives who have the same gender and age structure and/or the same regional distribution. This is a "counterfactual" question, in the sense that it refers to comparing hypothetical populations that are identical in a set of observable characteristics. We can answer such questions by including a vector of observable variables *X* [Sex, Age, Age^2] in our regression. In this case, the coefficient measures the difference in the benefits receipt rate (probability of living in social housing [?]) between immigrants and natives who are identical with respect to the variables included in *X*. Our *empirical* analysis addresses both factual and counterfactual questions. (**p.10**)
> [My emphasis.]

The 'empirical' here is a reminder to ask "Where is the evidence that (4) and its submodel (1) have any rational basis as a mathematical description of reality?" Formula (4) is strictly additive with no interaction terms that might do something to represent reality. There are many areas of science where functions of several variables can be well-approximated by such additive formulae, but only under strict conditions that cannot be seen to hold in the present context. The onus of proof must then be on those who present us with this simply additive formula. It does not encourage faith in (4) that the Discussion says that there may have been another explanatory factor ('education') in the estimation of the coefficient −0.084 (if the factor had been added in full, there would have been at least two more dummy variables in the already-empirical (4)). The halving of β^* to β^{**} is not a small adjustment and it goes half-way to zero, but Cream has no doubt about what it signifies :

> Recent immigrants are 45% (18 percentage points) less likely to receive state benefits or tax credits. These differences are partly explainable by immigrants' more favourable age-gender composition. However, even when compared to natives with the same age, gender composition, and education, recent immigrants are still 21% less likely than natives to receive benefits. (pp.28-29)
> [The arithmetic is 45% is 0.178/0.391 and 21% is 0.084/0.391.]

Statistical theory has important, but unavoidably technical, things to say that are relevant to the inferential significance of the coefficients, standard errors and asterisks in Table 3:

(i) There is a proof (on weak assumptions and open to inspection by any reader) that, when estimated by OLS, the standard error of β^* in Model 1 (for 4+5 Recent immigrants) is less than 0.001—with which Table 3's 0.001 scrapes agreement. But this upper bound corresponds to the total-ignorance model that uses the 1/0 outcome of the toss of a fair coin to fit D! The corresponding value for 'robust' estimation is unlikely to be much larger, simply because a binomial variance is only weakly heteroscedastic for the binomial probabilities corresponding to the observed proportions $p_q(D=1|I=0)$ and $p_q(D=1|I=0)$, which are likely to be in the range 0.1 to 0.9. The fact that the Table 3 value is no lower than 0.001 (i.e. 0.0005) tells us that Model 1 can be said to be a poor fit of the 3.5m values of D. But that, in itself, is not a reflection on the choice of model—the poverty is only to be expected. Even the relaxation of the model to $D \approx \beta_q I + \tau_q$ (which has 88 parameters (one for each Immigrant-status x Year-quarter category) would also be a poor fit of D, although it could possibly (who knows?) be a reasonable fit of the aggregates of D in $p_q(D=1|I=0)$ and $p_q(D=1|I=0)$. To get a really good fit to D, you would have to be able to predict the variation of individuals' values of D *within* the Immigrant-status x Year-quarter categories—and that would need much more information than is to be found in LFS records).

(ii) A more worrying reflection on the undocumented fit of the Cream models is that (as my Appendix establishes, without assumptions) the largest coefficient, 0.178, accounts for less than 3.5% of the variation of D within year-quarters (in the squared residual sense) that the inclusion of β in Model 1 is trying to explain. (If the 0.001 were the 0.0010 of the theory in (i) and not the conservative 0.0005, the 3.5% would be 1%.) In other words, the figures in Table 3 tell us what we would have been told by a low value of the R^2 that Cream does not report—namely, that there is (as is to be expected) a large *terra incognita* beyond what Model 1 manages to explain of the variation of D (even with the β). With three (or is it five?) more parameters, Model 2 will be explaining more of the total variation than Model 1, but until Cream reveals the value of R^2 we will never know how much more.

(iii) The most important comment, however, on what Cream infers from Table 3 is analogous to the grounds on which statisticians reject the Nuffield Trust formula for NHS England's funding of Clinical Commissioning Groups (see www.civitas.org.uk/pdf/stoneseries). In both areas, a crudely-additive empirical formula is imposed on the data and the inevitable statistical significances of the estimated coefficients are taken to justify an implementation of the formula in the NHS case and a rebuttal of a popular prejudice in Cream. No acknowledgement is given of the fact that:

(a) huge sample sizes (5 million in one case and 3.5 million in the other) can give statistical significance to very small coefficients (e.g. the 0.004 for Non-EEA Model 5)

(b) a quite different empirical formula may explain a larger (but still small) fraction of *terra incognita*—and give different values of the sizeable coefficients that determine CCG allocations in one case and naïve inferences about 'probability gaps' in the other.

Conclusion

Cream set out to provide no more than *substantive evidence on immigrants' fiscal contribution* but the authors' own estimation of their achievement is that they have done better than that—claiming that they have provided a *clean picture*, namely:

> *precise* estimates for each year since 1995 (2001 for recent immigrants) on both the overall expenditure on the respective immigrant populations and the revenues they have produced in comparison to native born workers. (p.3) [my emphasis]

If any honest statistician had made the same painstaking but assumption-based calculations, the last word he/she would have used to describe the estimates is 'precise' (unless exhaustion had affected judgement). Most of the underlying crude assumptions that the all-embracing approach has been obliged to make have not been subject to sensitivity tests that might been made if the study had not been so obviously driven to make the case it claims to have made. As we have seen, even the tabulation of results lacked the statistical straightforwardness necessary for neutrality of presentation.

There is less justification for questioning the neutrality of the econometric manipulation of LFS data that Cream has adduced to support its fiscal accountancy—for the simple reason that such manipulation of data by purely empirical models is what econometricians are trained to do without question. I have, however, tried to give good reason why readers should maintain a degree of scepticism about the results of the models that Cream has employed to boost its case.

Appendix

(a) Preparatory algebra for deriving expression (1) for β*

The 'linear probability model' OLS fitting of formula (2) finds values of β and $\tau_1, \ldots, \tau_{44}$ that minimize the total (from 2001 to 2011) of a squared measure of how badly the formula fits the observed values (1 or 0) of D. For the four categories of individual, the possible values of (observed − fitted)^2 are:

	I = 1	I = 0
D = 1	$(1-\beta-\tau_q)^2$	$(1-\tau_q)^2$
D = 0	$(0-\beta-\tau_q)^2$	$(0-\tau_q)^2$

The total measure is therefore the sum of the 44 quantities

$$a_q(1-\beta-\tau_q)^2 \; + \; b_q(1-\tau_q)^2 \; + \; c_q(\beta+\tau_q)^2 \; + \; d_q\tau_q{}^2 \; .$$

High-school pupils or readers, who know enough differential calculus to minimize a quadratic function of 45 parameters and enough algebra to solve 45 simply-structured simultaneous equations, should have no difficulty in verifying equation (1) for minimizing β*.

(b) A truly-significant message about the asterisks and standard errors of Table 3

A very large and astronomically significant t-value such as 356 = 0.178/0.0005 (taking the quoted standard error to be as low as 0.0005) can easily arise when, as here, the sample size (the N of 3.5m) is also very large. For OLS estimation, the fractional reduction of a Residual Sum of Squares by adding one more parameter (such as β in Models 1 and 2) to a model with p parameters is $t^2/(t^2+N-p-1)$; in Model 1, p is 44 and, in Model 2, p is 47, say[5]). For the coefficient −0.178, the fraction is only 0.0349. For the coefficient −0.084 with t-value 168 = 0.084/0.0005, the fraction is only 0.008.

[5] Galbraith, J. and Stone, M. (2011), The abuse of regression in the National Health Service allocation formulae: response to the Department of Health's 2007 'resource allocation research paper'. Journal of the Royal Statistical Society: Series A (Statistics in Society), 174: 517–528. doi: 10.1111/j.1467-985X.2010.00700.x