UNIVERSITY OF OXFORD

Oxford Economic and Social History Working Papers

Number 204, December 2022

The Black Death and the Origin of the European Marriage Pattern

Jeremy Edwards and Sheilagh Ogilvie

The Black Death and the Origin of the European Marriage Pattern

Jeremy Edwards
Faculty of Economics
University of Cambridge
Sidgwick Avenue
Cambridge
CB3 9DD
jsse12@yahoo.co.uk

Sheilagh Ogilvie
All Souls College
University of Oxford
Oxford
OX1 4AL
sheilagh.ogilvie@all-souls.ox.ac.uk

Abstract: This paper evaluates criticisms of our view that there is no evidence of the Black Death having caused the European Marriage Pattern. The attempt by Nico Voigtländer and Hans-Joachim Voth to rebut our argument fails completely. Their claim that we distort the historical evidence is entirely without foundation. They do not engage with the fact that historical demographers are widely divided on when this marriage pattern emerged and that the data are too fragile for any definitive conclusions about the period before c. 1540. They sidestep the logic of their own model, refusing to acknowledge that the factual inaccuracy of one key assumption makes the model completely inapplicable to demographic behaviour in England after the Black Death. They repeatedly refer to evidence on demographic behaviour centuries later than the Black Death with no attempt to explain how it could be relevant to the aftermath of that pandemic. This weakness also applies to the econometric evidence they adduce, but that evidence is further vitiated by the invalidity of the instrumental variables they use.

JEL classifications: E02, J12, J13, N13, N33.

1. Introduction

The Black Death freed England from the Malthusian trap, it has been argued, by increasing wages in pastoral agriculture, thereby delaying female marriage, restricting fertility, and giving rise to the European Marriage Pattern (Voigtländer and Voth 2013, henceforth VV 2013).[1] We showed in a recent article that this argument does not hold (Edwards and Ogilvie 2022). We began by demonstrating that there is no scholarly consensus that the European Marriage Pattern (EMP) in general, or late female marriage in particular, emerged in rural England after the Black Death. We continued by surveying historical evidence showing that in medieval England women wanting to do pastoral work did not have to remain unmarried, so enjoying higher pastoral earnings did not necessitate later marriage. We then scrutinized the empirical relationship between pastoralism and female marriage age in England, and showed that it provides no support for the argument that the growth of pastoral agriculture after the Black Death caused women to delay marriage. We concluded that there is no evidence that fertility restriction in medieval England was exogenously triggered by the Black Death.

In a response to our article, Voigtländer and Voth (2022) (henceforth VV 2022) criticise our arguments and seek to defend their original claim that the Black Death caused economic development by "inventing" fertility restriction. However, their response fails to address our criticisms. In what follows, we show that VV's reply (VV 2022) suffers from the same flaws as their original article (VV 2013).

[1] The European Marriage Pattern is the name given to a demographic system involving women marrying late or not at all, predominantly nuclear-family households, and the circulation of young people between households before marriage as live-in servants. Fertility consequently lay below the biological maximum and, because marriage involved forming a new household, responded to economic conditions. The concept of the European Marriage Pattern was originally proposed by Hajnal (1965, 1982, 1983); for a survey of the literature, see Dennison and Ogilvie (2014, 2016).

The basic premise of their argument – that fertility restriction originated in the aftermath of the Black Death – is not accepted by scholars in the field. VV's attempted rebuttal provides no evidence that nuptiality or fertility declined in rural England in the aftermath of the Black Death. No such evidence exists, which is why the idea is not accepted in the scholarship on medieval demography.

A crucial assumption of the theoretical model underlying VV's claim is that women had to stay unmarried in order to earn high post-pandemic pastoral wages because the only way women could work in pastoral agriculture was to be employed by landlords as unmarried servants. This assumption is factually inaccurate, as we showed in Edwards and Ogilvie (2022). For one thing, married women could also earn income doing pastoral work for landlords after the Black Death. More importantly, peasant households engaged in market-oriented pastoral production both before and after the pandemic. Peasant women could therefore benefit from the rise in pastoral earnings without having to refrain from marriage. VV state several times in their reply that their argument is not affected by this evidence. But repeating a false claim several times does not make it true. Without the assumption that pastoral work could only be done by unmarried women, VV's model is unable to explain why the higher earnings in pastoral agriculture after the Black Death should have led to later female marriage and lower fertility.

The empirical analysis carried out by VV also suffers from a number of serious flaws. These include using an inappropriate proxy for female celibacy, employing invalid instrumental variables, and relying on evidence of demographic behaviour centuries later than the Black Death. In their reply, VV seek to defend their proxy for female celibacy and their use of instrumental variables that violate the exclusion restriction, but without success. They also repeatedly seek to shore up their argument using evidence from much later historical eras, without explaining why this evidence might be relevant to assessing the effects of the

Black Death. In putting forward a claim that economic development was triggered by a specific historical pandemic, this ahistorical approach is a fatal weakness.

2. The Historical Evidence

The claim that the Black Death caused the emergence of the EMP in England is inconsistent with what is known about English demographic and economic behaviour before and after the Black Death. Two components of the VV argument in particular are not supported by the historical evidence.

One is the claim that there is a consensus among demographic historians that the EMP only became fully developed after the Black Death. Scholars hold widely divergent views about when exactly the EMP developed, but in so far as there is a consensus it is that the EMP did *not* emerge in rural England soon after the Black Death.

The other component of the VV argument that is rejected by the historical evidence is the mechanism which purportedly generated lower fertility as a result of the shift to pastoral agriculture after the Black Death: the assumption that women could do pastoral work only as servants employed by landlords, which required them to remain unmarried. The historical evidence shows that women did not have to be unmarried to work in pastoral agriculture in fourteenth- and fifteenth-century England. Peasant wives could do market-oriented pastoral work in peasant households which, although individually small, collectively produced a significant amount of pastoral output nationally. Married women could also do pastoral work for landlords as labourers who were paid a daily wage.

VV (2022) claim that our interpretation of the historical evidence is flawed. In this section we discuss the criticisms VV level at our interpretation, and show that none of them can be sustained.

2.1 Distortions?

VV state on p. 1248 of their reply that we “frequently distort the meaning and content of historical scholarship that [we] cite”, on p. 1252 that “distortions abound”, and on p. 1254 that we “cite most sources in a highly selective manner, to the point of misrepresenting the views and historical assessment of the authors involved”. These allegations would require a great deal of supporting evidence to be justifiable, but VV provide only three examples of our supposedly distorting the historical evidence we cite. None of these three examples stand up to scrutiny.

The first example VV provide concerns our reference to Watkins (1989, p. 19). VV state that “Watkins (1989, p. 19) says nothing about the alleged ‘shift from landlord to peasant production’ ”. VV’s statement is false.

We refer to Watkins (1989) as follows (on p. 1233 of our paper):

> One way peasants expanded pastoral production was by leasing pastures from landlords (Dyer, 1981, pp. 4–5; Watkins, 1989, p. 18). Such leases already occurred before 1350 but proliferated rapidly thereafter, and by c. 1450 most landlords had abandoned direct exploitation of their demesnes (Lomas, 1978, pp. 339–40, 345, 352; Campbell, 2000, pp. 3, 58–60; Dodds, 2008, p. 77). This enabled many more peasants to get access to grazing and begin raising livestock (Bailey, 1989, pp. 257–8; Watkins, 1989, p. 19; Campbell, 1992, p. 113; Bailey, 2007, pp. 220–1; Hare, 2011, pp. 78–9, 101–5).

The relevant parts of p. 19 of Watkins (1989) are:

> The Baillys of Middleton were another peasant family who were able to acquire considerable lands and wealth through cattle grazing ... by 1420 [Thomas Bailly] was leasing considerable parts of the demesne

and

> Before the detailed probate inventories starting in the sixteenth century it is extremely difficult to draw any precise conclusions about the livestock keeping of ‘ordinary’ peasants. Although the evidence is limited, it has been generally accepted that in the post-plague period livestock ownership became more widespread and disseminated throughout peasant society.

Watkins (1989) therefore provides on p. 19 of his article a detailed example of a peasant family becoming wealthy after the Black Death by leasing pastoral demesne land, and a clear statement that it is generally accepted that in this period livestock ownership became more widespread among peasants. The latter statement is the standard view in the historical literature, and it is supported by many other references which are cited on p. 1233 of our paper. But it is also a view with which Watkins (1989) agrees, as is shown by his statement on p. 19 that "it has been generally accepted that in the post-plague period livestock ownership became more widespread and disseminated throughout peasant society". If it is objected that Watkins' agreement with this view on p. 19 is only implicit, let us quote from p. 17 of his article:

> Studies throughout the country have shown the ability of elements within rural society to take advantage of the landlords' withdrawal from a market-orientated productive economy by taking up demesne leases. As elsewhere the wide social range of lessees, including clergy, gentry, merchants, and peasants is apparent in Warwickshire.

The argument we make in the quotation above from p. 1233 of our paper is clearly one with which Watkins (1989) agrees. VV's claim to the contrary is simply false.

On pp. 1251-2 of their reply, VV claim that some statements they quote from Watkins (1989, pp. 15-16) show that he

> argues that landlords engaged in more pastoral production, hiring predominantly women ... after the Black Death. This is exactly the shift ... that creates the basis for fertility limitation in our argument. Watkins also points out that for smaller producers further down the social scale, the evidence for a shift to pastoralism is much less clear.

However, the statements from Watkins (1989) that VV quote have been extracted from more than one whole page of Watkins' article and stitched together into a "quotation" conveying a very different meaning from that which emerges when these statements are read in the context of Watkins' original text that VV omitted.

The statements by Watkins (1989) that “The attractions of cattle for landlords were considerable ... They cost little to maintain ... Neither were labour costs very great[2] ... More labour intensive was dairying, a task usually undertaken by women both in processing and retailing” come at the end of a section of Watkins’ article in which he discusses the difference between the livestock holdings of landlords in southern Warwickshire, where sheep were predominant, and the Forest of Arden in northern Warwickshire, in which the majority of livestock were cattle. Watkins’ statement about the attractions of cattle thus refers to the attractions for landlords in the Forest of Arden of cows and beef animals relative to other domesticated animals such as sheep; it does not refer to a general increase in pastoral production by landlords. Watkins’ observation that dairying was usually undertaken by women is not evidence that landlords expanded pastoral production after the Black Death and did so primarily by employing more women. At most, it might suggest that the attractiveness of cattle relative to sheep meant that landlords hired more women to do dairying. Watkins (1989) says nothing about the marital status of the women who did dairying, so this statement provides no evidence in support of VV’s claim that women had to be unmarried to work in pastoral agriculture.

Watkins’ statement that “Below the level of the nobility and the religious houses the evidence of pastoral activity is less plentiful” occurs roughly 450 words later in his article and in a different section from the other statements quoted by VV, and thus should not be linked with them as if it were part of a contiguous passage of text. In any case, the fact that evidence of pastoral activity below the level of the nobility and religious houses is less plentiful does not mean that it does not exist: indeed, in that part of the text of p. 16 of Watkins’ article which VV omit when quoting from it there are two references to landlord

[2] VV do not quote Watkins (1989) accurately here: according to them he said “nor were labour costs very high”.

demesnes purchasing cattle from peasants, showing that peasants were engaged in pastoral agriculture.

VV's second example of our purported distortion of the historical evidence concerns Dyer (2005). VV state on p. 1252 of their reply that we cite Dyer (2005) "extensively to suggest that peasants shifted toward pastoral farming more than landlords did" and allege that this misrepresents Dyer's view. To support this allegation, VV quote from Dyer (2005, p. 72) as follows:

> A gazier [sic], Andrew Bate, in 1466 was keeping so many cattle that tenants were driven from one end of Dengemarsh. Landowners like the farmers Thomas Roby and Thomas Holderness were systematically buying up smallholdings between 1480 and 1520, and the resident population was falling as the land came into the hands of gentry. . . .

VV claim on pp. 1252-3 of their reply that this quotation

> is fully in line with VV's argument – more large-scale pastoral farming after 1350, through evicting peasants, buying up land, and enclosing common lands. Evidently, enclosure is only one part of this process, that is putting fences around common land. As the text by Dyer suggests, evicting peasants and buying up smallholdings were much more important, quantitatively and socially.

But on the very page of Dyer (2005) from which VV quote, p. 72, Dyer also states that, "Against these high-profile activities of the elite we need to place the significant and cumulatively larger-scale engrossing, enclosure, and conversion to pasture by the peasants themselves". Dyer therefore clearly states that peasants moved to pastoral agriculture on a larger scale than did landlords. VV's argument here is simply not made in good faith, as it does not include relevant material on the precise page in Dyer (2005) from which they quote. The allegation they make on pp. 1252-3 of their reply quoted above is false. There is no distortion of Dyer (2005) in our paper.

VV's third example of supposed distortion on our part is their contention on p. 1254 of their reply that we refer to Poos (1991) "to support [our] claim that women working in husbandry were often married and that there is no link between the use of live-in servant women in livestock farming and late marriage." However, we do not refer to Poos (1991) in the manner VV contend.

Our view that there were plenty of opportunities for women to marry and nonetheless benefit from high pastoral earnings is primarily based on the evidence discussed in our paper showing that peasant households were actively engaged in market-oriented pastoral agriculture both before and after the Black Death. We do not claim that women employed by landlords in pastoral agriculture were "often married", and it is misleading of VV to imply that we do. We simply point out that there is evidence showing it was possible for married women to be employed in pastoral agriculture by landlords. Even if the women employed by landlords in pastoral agriculture were never married, VV's argument would still not apply, because of the evidence that market-oriented pastoral production was undertaken by peasant households, including by married peasant housewives.

However, some of the female wage-labourers employed in pastoral agriculture by landlords were indeed recorded as being married. It is in that context that we refer to the evidence presented in Poos (1991). We refer to Poos (1991) twice for evidence that married female wage-labourers existed, and once for evidence that a married female wage-labourer was explicitly employed for a pastoral task. We are careful to point out how sparse is the evidence of the tasks performed by married female wage-labourers before c. 1480. However, wage-labouring was not the only way in which married women were employed in pastoral agriculture by landlords. We provide evidence, from sources other than Poos (1991), of married women working for landlords in vaccary-keeping familial teams and as managers of demesne dairies (see pp. 1235-6 of our paper). We make no claim about the frequency with which the women employed by landlords in pastoral agriculture were married, and VV's claim that we cite Poos (1991) to the effect that such women were "often married" is false.

VV's contention that we refer to Poos (1991) in support of a claim that there is no link between life-cycle servanthood and late marriage is also false. As has just been explained, our references to Poos (1991) are solely concerned with the evidence he presents which demonstrates that married female wage-labourers existed in the medieval English rural economy. We do not use our references to Poos (1991) to say anything about life-cycle servanthood and late marriage. Furthermore, we do not claim that there is no link between life-cycle servanthood and late marriage. Rather, we point out that the evidence shows that medieval English women did not have to be unmarried servants in order to work in pastoral agriculture, and that this must be taken into account in any explanation of the emergence of the EMP. VV's analysis does not do so.

We agree with VV that Poos (1991) has a different view of the late medieval agricultural economy in England than we do, but we do not suggest otherwise in our references to his work. We quote Poos (1991) only to refer to the examples he provides from the primary sources which testify to the existence of married female wage-labourers. The fact that Poos (1991) has a different view to ours does not mean that we have distorted Poos (1991) by using some of the evidence he himself provides of married female wage-labourers. In any case, Poos (1991) provides no actual evidence to support his view, and is not the final word on the nature of the late medieval agricultural economy. Our view is based on a great deal of research that has appeared in the 30 or so years since Poos (1991) was published.

In summary, VV provide just three examples to justify their allegation that we distort the historical evidence to which we refer. As this section has shown, in each of these three cases our references to the historical scholarship were meticulously accurate and involved no distortion. We completely reject VV's allegations.

2.2 When Did the EMP Emerge?

In order for it to be possible to argue that the Black Death caused the EMP to emerge in England, it is necessary to establish two things: first, that the EMP did not exist in England before c. 1350; and second, that the EMP is unquestionably observable there in the aftermath of the Black Death. The historical evidence, however, does not allow any definitive statement to be made about when precisely the EMP emerged. In our article, we noted (on p. 1231) that a number of historical demographers (Richard Smith, Judith Bennett, H. E. Hallam, and the originator of the EMP concept, John Hajnal, in his later publications) think that the EMP existed in England in the thirteenth century, and thus before the Black Death. However, a number of other economic historians of medieval England (John Hatcher, Mark Bailey, Mavis Mate, and Zvi Razi) think that the EMP did not exist in England until the early sixteenth century, at least 150 years after the Black Death. The absence of any definite evidence of when the EMP emerged in England means that VV's argument about the causal role of the Black Death rests on very shaky foundations.

VV claim that "there is little to suggest that Europeans married late (or little) prior to the 14th century" (VV 2022, p. 1249). But this is not true. VV simply ignore the arguments of the scholars cited on p. 1231 of our article (Smith, Bennett, Hallam, and Hajnal himself) who regard the EMP as having been present in England in the thirteenth century. Ignoring an argument does not amount to rebutting it, as would be necessary for VV's claim to be correct.

In response to our pointing out that "no historical demographer argues that the EMP emerged in the *rural* economy immediately after the Black Death", VV state that they did not argue that the EMP emerged immediately after the Black Death, so their argument cannot be criticised on this ground (VV 2022, p. 1249). This is not true. In their original article, VV (2013, p. 2227) write, "[a]s early as the fourteenth century, a 'European Marriage Pattern'

(Hajnal 1965) had emerged which combined late marriage for women with a significant share of women never marrying". Furthermore, one piece of evidence VV adduce in support of their argument is a claim that more pastoral areas of England had higher proportions of unmarried women in 1381 – only thirty years after the Black Death.

VV claim that their theoretical model actually implies that the EMP should not have emerged in the aftermath of the Black Death. On p. 1249 of their reply, they write, "Had EMP emerged immediately after the Black Death, it would be a problem for our analysis – particularly high land-labor ratios predict early marriage and less female labor in pastoral farming in our model (see Fig. 4 in VV [2013])". But VV seem to have difficulty understanding their own model. Figure 4 of VV (2013) shows average female labour supply in pastoral agriculture increasing, and average fertility decreasing,[3] as the land-labour ratio in pastoral agriculture increases, up to some intermediate value of this land-labour ratio. Beyond this value, further increases in the land-labour ratio result in decreases in average female pastoral labour supply and increases in average fertility. However, at the highest value of the land-labour ratio in pastoral agriculture shown in Figure 4 of VV (2013), average female labour supply in pastoral agriculture is higher, and average fertility is lower, than is the case for the smallest 20 per cent of values of this land-labour ratio. VV's statement above is thus not true in general. In their model, it is perfectly possible for an increase in the land-labour ratio from very low to very high values to result in a decrease in fertility.

VV's model thus does provide for the EMP emerging immediately after the Black Death, and that is the claim VV make in their empirical analysis. However, it is not the consensus of the scholarship on medieval English demography.

[3] A decrease in average fertility corresponds to an increase in average marriage age in the VV model.

2.3 Life-Cycle Servanthood

An essential component of VV's explanation of the putative emergence of the EMP after the Black Death is the prevalence of life-cycle servanthood for females in England at the time of the Black Death and in its immediate aftermath, which, in their view, required women who wished to take advantage of the greater opportunities in pastoral agriculture to remain unmarried. Life-cycle servanthood – young, unmarried people of both sexes living and working in their employers' households for several months or years between leaving their parents' home and getting married – was an integral part of the EMP (Hajnal 1982, pp. 470-6). The problem with VV's argument is that one key feature of the EMP – life-cycle servanthood – is being used to explain the emergence of two other key features of the EMP – fertility restriction and the responsiveness of fertility to economic conditions. To put it another way, VV's assumption that women could only work in pastoral agriculture after the Black Death as unmarried servants implies that one key feature of the EMP already existed c. 1350, in which case it seems likely that the other key features also existed c. 1350 and thus cannot have been caused by the Black Death. The VV argument appears to be internally inconsistent.

Rather than attempting to explain how it might be possible for the life-cycle servanthood component of the EMP to have existed before the fertility restriction components of the EMP, VV respond to this point by advancing an idiosyncratic definition of life-cycle servanthood. In their view, life-cycle servanthood involves the combination of pastoral agriculture being performed to an important degree by large landowners, women doing pastoral work for such landowners typically having to remain unmarried, and pastoral agriculture becoming more important after 1350 (VV 2022, p. 1249). On the basis of this private and idiosyncratic definition of life-cycle servanthood, VV then assert that "[s]ince pastoral production increased significantly after the Black Death, so did life-cycle servanthood. In other words, we do not require life-cycle servanthood to have been a prominent feature of the pre-1350 economy" (VV 2022, pp. 1249-50).

This assertion makes no sense. There is no actual evidence that servanthood in husbandry by unmarried women increased after the Black Death. A number of scholars view the period from c. 1350 to c. 1500 as one in which rural women married early and earned relatively high wages from casual labouring, so that rural life-cycle service, if it existed to any significant extent at all in this period, became less important (Bailey 1996; Mate 1998). We noted in our article (p. 1235) that Humphries and Weisdorf (2015) show that females could earn much more as day-labourers than as servants during this period, which supports this view. VV allege that our interpretation of the Humphries-Weisdorf (2015) findings about the day-wages earned by female labourers and the daily remuneration implied by annual contracts for female servants is flawed, because no account is taken of the in-kind remuneration received by servants (VV 2022, p. 1254 fn 3). This is false: Humphries and Weisdorf devote a great deal of attention to valuing payments in kind (see pp. 413-7 of their article), with the result that their conclusions about the relative remuneration of female servants and labourers take full account of payments in kind.

The increase in pastoral production after c. 1350 is undisputed, but this does not necessarily imply that there was any increase in the number of unmarried female servants employed in pastoral agriculture, since married women were able to do market-oriented pastoral work either as day-labourers or as peasant housewives. In any case, nothing except obfuscation is achieved by VV adopting an idiosyncratic definition of life-cycle servanthood which bears no relation to the usual definition in the literature. The standard definition of life-cycle servanthood does not involve pastoral production or large landowners as integral characteristics (Hajnal 1982; Goldberg 1986; Whittle 2017). In the words of Laslett (1988, p.

57), under the institution of life-cycle service in pre-industrial England “Servants are found in households of all kinds and occupations and even labourers’ households can have servants ... appreciable numbers of servants in the countryside as well as in the towns come from families on the same social level as their employers”. VV’s reliance on the existence of life-cycle servanthood to explain the emergence of fertility restriction and the responsiveness of fertility to economic circumstances means that they use one key feature of the EMP to explain the existence of two others. They cannot resolve this problem with their argument by adopting their own private definition of life-cycle servanthood.

VV use their theoretical model, a central assumption of which is that women can only work in pastoral agriculture as unmarried servants, to put forward an explanation of why the EMP did not emerge after the Black Death in Eastern Europe and China. However, historical demographers do not find evidence that life-cycle servanthood existed in Eastern Europe and China in the medieval period. As a result, the VV model is not applicable to these countries. Therefore it cannot explain why the Black Death did not lead to the emergence of the EMP in Eastern Europe and China. In response to this point, VV again resort to their idiosyncratic definition of life-cycle servanthood which portrays it as integrally connected with pastoral agriculture. They claim that in Eastern Europe and China arable productivity was so high that it prevented a shift to large-scale pastoral agriculture, and so, according to their private definition, there was no life-cycle servanthood (VV 2022, p. 1250). But, to repeat, life-cycle servanthood is not integrally connected with pastoral agriculture in actual historical fact (Hajnal 1982; Goldberg 1986; Laslett 1988; Whittle 2017). So the VV model cannot adduce an absence of pastoral agriculture in Eastern Europe and China to explain the absence of life-cycle servanthood there. In any case, the VV model cannot be used to explain why life-cycle servanthood exists: the existence of life-cycle servanthood for women who wish to work in pastoral agriculture is an assumption, not a result, of that model.

VV insist that “[t]here can be no serious question that male and female teenagers worked as servants in agriculture on a large scale, and that they were overwhelmingly celibate” (VV 2022, p. 1248). They claim that they provide “[t]he simple facts about arable versus pastoral farming, and the use of female labor” (VV 2022, p. 1251) in England after the Black Death. However, the “simple facts” VV provide date from the mid-sixteenth century and later, and so are not relevant to English agriculture in the 150 years or so after the Black Death. VV refer to Kussmaul (1981) in support of their claim that unmarried servants played a major role in agricultural labour and were “overwhelmingly female”. However, as we pointed out on p. 1232 of our article, Kussmaul only has evidence for the period 1538-1840 and states explicitly that any claim about the importance of life-cycle servanthood in the fifteenth century is speculative. Furthermore, VV report Kussmaul’s findings incorrectly. VV cite Kussmaul (1981, p. 4) as the source for their claim that “[s]ervants in agriculture were also overwhelmingly female – 213 females per 100 men” (VV 2022, p. 1251). But this is exactly the wrong way round. In fact, Kussmaul (1981, p. 4) reports that, according to the 1851 census, there were 213 males to 100 females in farm service.

VV also seek to buttress their claims (VV 2022, pp. 1251-2) by referring to a figure derived from Allen’s (1991) reworking of English agricultural data from the late 1760s. Data from a period more than 400 years after the Black Death are not relevant for understanding the role of female labour in arable and pastoral agriculture in the aftermath of the Black Death. In any case, the data in Figure 1 of VV (2022) show only that proportionally more women worked on pastoral than arable farms, and that this difference was greater on larger farms. But labour provided by women is not the same as labour provided by unmarried female servants: as shown by the historical evidence discussed in our paper, agricultural labour was also provided by married women. VV’s Figure 1 provides no evidence demonstrating that unmarried female servant labour became more important in English agriculture in the aftermath of the Black Death, and the claim that it did so is explicitly disputed by several economic historians of medieval England (e.g. Bailey 1996; Mate 1998, 1999a).

2.4 Peasants in Pastoral Agriculture

Our article provides evidence that much pastoral output in England before and after the Black Death was produced by peasants, so peasant women did not have to delay marriage and work for landlords in order to gain from rising pastoral earnings. VV attempt to cast doubt on this evidence by claiming (VV 2022, p. 1251) that a defining characteristic of “peasant” is that a peasant does not hire others to work. This is simply false. The definition of “peasant” that VV quote from the Cambridge Dictionary makes absolutely no reference to whether a peasant hires others to work. Furthermore, economic and social historians of medieval England have shown that the more substantial peasants supplemented family labour by employing servants on annual contracts and labourers working for wages on a part-time basis (Hilton 1975, pp. 30-5; Dyer 1981, p. 24; Hilton 1985, p. 263; Dyer 2005, pp. 56, 129). The accepted use of the term “peasant” by scholars of medieval English economic and social history does not imply the absence of hired labour. VV’s claim (VV 2022, p. 1253) that their definition of “peasant” is in line with standard practice is false.

VV regard any pastoral producer who hires labour from outside the household as a ”landlord” for the purposes of their model (VV 2022, p. 1251). This means that the pastoral producers in medieval England who employed outside labour but rented the land on which they worked are ”landlords” as far as VV are concerned. There is no accepted use use of the term “landlord” in English which applies to someone who rents rather than owns land. VV’s attempt to claim that that our criticisms of the assumptions underlying their analysis reflect our failure to understand the meaning of “peasant” and “landlord” has a distinct flavour of what Paul Romer has labelled “mathiness”: the use of a mixture of words and symbols that leaves ample room for slippage between statements in the languages of words as opposed to symbols (Romer 2015). VV’s model is apparently to be understood only by adopting meanings of “peasant” and “landlord” that bear no relation to either standard or scholarly usage. The same point can be made about VV’s adoption of their own private meaning of “life-cycle servanthood”, discussed in the previous sub-section.

VV allege that, in referring to the estimates by Masschaele (1997) of landlord and peasant wool production in early-fourteenth-century England, we “suggest a consensus where there is none among medievalists” (VV 2022, p. 1253). VV do not mention that in our discussion of the relative importance of landlords and peasants in wool production we also refer to Bridbury (1977) and Campbell (2000), who reach similar conclusions to those of Masschaele (1997) about the importance of peasant wool production. VV do not provide references to any scholar who takes a different view to that of Bridbury, Masschaele, and Campbell about the important role of peasants in the production of wool before the Black Death, so their claim that there is no consensus is conspicuously lacking in evidence.

VV state that Masschaele’s estimates are at variance with the evidence on medieval royal taxation in Maddicott (1975). But this is not correct. Masschaele does base his estimate of the share of *total agricultural output* produced by peasants on evidence from lay subsidy rolls and royal purveyance accounts, and Maddicott’s work does raise questions about the use of subsidy rolls and purveyance accounts for such a purpose (Mate 1999b). However, Masschaele’s estimate of the importance of peasants in *wool production* is based on evidence of the quantity of wool exported in the early fourteenth century derived from customs accounts combined with information about the wool production by ecclesiastical and secular estates. There is nothing in Maddicott’s work which casts doubt on the sources used by

Masschaele to estimate the importance of peasants in wool production, and hence in pastoral activity.

VV's assumption that peasants could not engage in pastoral agriculture is clearly untenable. However, similar results to those of the VV theoretical model would be obtained if instead it was assumed that pastoral production required a minimum amount of land (VV 2013, p. 2238 fn. 27). We discussed this alternative assumption on pp. 14-15 of the Appendix to our article, and showed that the historical evidence makes it untenable as well. In their reply, VV quote part of our discussion of this point, and use their private and idiosyncratic definition of "peasant" to claim that the evidence we provide supports the alternative assumption of a minimum land requirement for pastoral production (VV 2022, p. 1253). This is false, as a full quotation of the discussion of this point in our Appendix would have made clear. VV only cite that part of our discussion in which we set out how the alternative assumption would lead to theoretical results similar to those obtained by adopting the assumption that peasants could not engage in pastoral agriculture at all. They omit our discussion of the evidence showing that, although peasants did indeed operate on a small scale individually, they were so numerous relative to landlords that they accounted for a very substantial share of pastoral production. In addition to the aggregative evidence provided by Bridbury, Masschaele, and Campbell already mentioned, we also referred to studies of the relative importance of peasant and landlord sheep, cattle, and poultry raising by Campbell (2000), Stone (2003), Bailey (2007), and Slavin (2010, 2015), all of which support the view that pastoral production by peasants was at least as important as that by landlords. It is not possible to make the VV model applicable to England in the aftermath of the Black Death by replacing the assumption that peasants could not engage in pastoral agriculture with the assumption that there was a minimum land requirement for such agriculture. The evidence we provide shows that peasants, in the standard meaning of the word rather than VV's private usage, contributed substantially to pastoral production, and thus that VV's theoretical model is inapplicable to fourteenth- and fifteenth-century England.

3. The VV Argument Again

The previous section established that VV's criticisms of our view of the historical evidence do not hold. It is unclear when the EMP emerged in England, but no scholar argues that it did so in rural areas immediately after the Black Death. VV claim that they never argued otherwise. However, as we showed in the previous section, in their original article they explicitly stated that the EMP emerged in the fourteenth century. They also purported to have evidence from 1381 supporting the mechanism by which the Black Death was supposed to have led to the EMP. In their reply, VV claim that the "process [by which the Black Death created an incentive to postpone marriage] evolved over centuries and that is how we frame it in [VV (2013)]" (VV 2022, p. 1249). Leaving aside the fact that VV did not so frame it in their 2013 article, it is completely unclear from VV's analysis how the mechanism by which they see the Black Death influencing marriage and fertility can have taken so long to operate. If the Black Death triggered a switch from arable to pastoral production which persisted over the next several centuries, it might be possible for VV to argue that there was a slow but persistent development of the EMP. However, according to Broadberry et al. (2015, pp. 60–1), cost and price incentives favoured a switch from arable to pastoral agriculture in England between about 1380 and 1500 and again after 1650, but for the 150 years from 1500 to 1650 arable farming was relatively more attractive. VV's argument implies, therefore, that female age at first marriage should have fallen and fertility should have increased during the entire sixteenth century and the first half of the seventeenth. How this can be made consistent with

VV's claim that the Black Death created an incentive for fertility restriction which evolved over centuries is obscure.

VV acknowledge in their reply that married women could work in pastoral agriculture in England before and after the Black Death, but repeatedly claim that this does not affect their argument. Thus, for example, they state that"[t]he fact that some pastoral production somewhere was performed by married women, and that some peasants kept cows and sheep is irrelevant for VV's argument" (VV 2022 p. 1250) and "[t]he feature of the VV model that matters is that a significant share of the labor necessary for livestock production was carried out by unmarried female labor" (VV 2022 p. 1253). These statements are false, as we shall now explain, drawing on our discussion in section 2.

In their reply, VV (2022, p. 1248) restate their argument about the origins of the EMP as follows:

> VV ... provide a unified explanation for Europe's strikingly late marriages. After a considerable share (perhaps up to one-half) of the population died, land-labor ratios rose, and wages surged. A good part of agricultural production switched from arable to pastoral farming. Women's labor is more compatible with pastoral activities ... Because of women's comparative advantage in livestock farming, this implied that as pastoral production surged, there was more work for women. When a good share of the extra labor required was provided through 'service in husbandry' (in particular, by young women working as servants in livestock production), fertility rates had to decline because farm service was 'a covenanted state of celibacy' (Davis, 1766) – not having children was part of the contract. In other words, the change in production after 1350 increased demand for female labor, which raised marriage ages and lowered fertility. VV summarize this argument in a mathematical model ...

As we pointed out in the previous section, there is no evidence that service in husbandry by unmarried women increased during the 150 years after the Black Death. VV's references to Kussmaul (1981) are not relevant, as she has no evidence for the period before 1538. Whether life-cycle servanthood existed to any significant extent in England in the 150 years after 1350 is a subject on which there is no scholarly consensus, as we explained in section 2.2 above. Rather, this question has given rise to lively debate among economic historians of medieval England. Some scholars do think that life-cycle servanthood was widespread in England before the Black Death, but they regard life-cycle servanthood as an integral part of the EMP, so that, if their view that servanthood existed before c. 1350 is accepted, it then follows that the EMP pre-dated the Black Death and cannot have been caused by it. Other scholars think life-cycle servanthood was not common in the English rural economy until the end of the fifteenth century. Even if there were substantial numbers of unmarried female servants working in pastoral agriculture in the mid-fourteenth century, several scholars think that the evidence suggests that unmarried female servant labour in agriculture became *less* important in the aftermath of the Black Death because working for day-wages as labourers became more attractive than working on annual contracts as live-in servants, as we explained in section 2.3 above.

The standard view held by medieval English economic historians is that the expansion of pastoral production after the Black Death involved peasants producing a greater share of pastoral output by taking up demesne leases from landlords, as we explained in section 2.1 above. This implies that much of the additional female labour in pastoral agriculture came from peasant wives working on peasant holdings. There is no evidence to support VV's view that "a good share" of the labour that moved into pastoral production after the Black Death took the form of unmarried female servants.

The role of formal models in economics is to establish the precise conditions under which an intuitive argument can be sustained. Such models inevitably make simplifying assumptions, but these assumptions are only acceptable provided that they are not absolutely essential for the results of the model, as we noted in Edwards and Ogilvie (2022, p. 1236). The assumption that women could only work in pastoral agriculture after the Black Death as unmarried servants is absolutely essential for the results of VV's model, as it ensures that women can only obtain the higher earnings in pastoral work by delaying marriage and thus having fewer children. However, in England after the Black Death women did not have to be unmarried to work in pastoral agriculture, so the choices involving pastoral work and marriage that women actually faced are not present in VV's model. If these choices were to be incorporated into the VV model, there would no longer be an obvious trade-off between higher pastoral earnings and having children, and the model would not yield the result that the higher pastoral earnings available to women after the Black Death led to a fall in fertility.

Suppose, for example, that the VV model, in which women obtain utility both from consumption of goods and from having children, was modified to allow peasants to produce pastoral output according to the same technology as landlords.[4] Then peasant women could benefit from the higher returns from pastoral work by undertaking pastoral work as wives in peasant households, without having to reduce the number of children they had. Increases in the land-labour ratio would no longer reduce fertility because there would be no trade-off between pastoral work and marriage, and VV's explanation of the effect of the Black Death on fertility would no longer apply.

The general point is that women did not have to be unmarried to work in pastoral agriculture, but VV's model depends critically on the assumption that they did. This assumption is not an acceptable simplification, and thus VV's model cannot provide the basis for an analysis of the effect of the Black Death on fertility in England. The one thing VV's model can do is to show why VV's claim that evidence that married women worked in pastoral agriculture does not affect their argument is false. The precise formulation of the VV argument that their model permits shows that it is critically dependent on married women being unable to do pastoral work.

[4] If this assumption does not hold, the analysis would be more complex, but there would still not be an unambiguous link between higher earnings for women in pastoral agriculture and lower fertility.

Our discussion of the inadequacy of the VV model so far has focused on its assumption that women had to remain unmarried in order to obtain the higher earnings in pastoral agriculture. However, the results of the VV model also depend critically on a second assumption, which is that women's preferences between consumption and children take a very special form, involving a basic-needs level of consumption. When income is below this basic-needs level, increases in income do not lead to additional children but solely to additional consumption. VV make no attempt either in their original article or in their reply to our article to explain why a theoretical analysis that is critically dependent on a very special and untestable assumption about preferences can cast any light on how the Black Death affected fertility.

Our view is that an explanation of how the Black Death caused the EMP which is based on a very special form of preferences contributes nothing to our understanding. It is not possible to bring any evidence to bear on the plausibility of this assumption, which is why we say no more about it, but this reinforces our point that the VV analysis is completely uninformative because of its crucial dependence on an assumption which is untestable.

4. The Econometric Evidence

In this section we discuss the comments VV make about our discussion of the econometric evidence they provide. Once again, VV's claims do not stand up to scrutiny.

4.1 Pastoralism and the Proportion of Unmarried Women in 1381

VV claim that their argument is supported by evidence from the 1377 and 1381 poll tax returns of a causal relationship between the proportion of pastoral land in a county and the proportion of unmarried women there. In our article we explain the fundamental flaw in VV's use of the proportional fall in the number of taxpayers between 1377 and 1381 as a proxy for the proportion of unmarried women in 1381 (Edwards and Ogilvie 2022, p. 1237 and Appendix Section A6). Using the work of Hilton (1975), Goldberg (1990), and Fenwick (1998), we show that the proportional fall in the number of taxpayers between 1377 and 1381 in fact comprises an unknown mixture of poor not-currently-married women, poor not-currently-married men, and poor married people of both sexes. Poor not-currently-married women are a mixed group that includes spinsters beyond child-bearing age and widowed females as well as never-married fertile women (the only group relevant for a test of VV's theory). Poor not-currently-married men, of course, consist only of never-married men and widowers, who are also irrelevant to testing VV's theory, as are poor married people. The proportional drop in taxpayers cannot, therefore, be interpreted as a proxy for the proportion of never-married fertile women.

VV's assertions to the contrary (VV 2022, pp. 1255-6) fail to acknowledge that married men, married women, widows, widowers, and spinsters beyond child-bearing age are all components of the proportional fall in taxpayers. This variable cannot, therefore, be interpreted as a measure of the proportion of never-married women of fertile age, which is what is needed to test their theory. VV try to shore up their choice of proxy by adducing a positive correlation between the proportional drop in taxpayers and the logarithm of the ratio of never-married men and widowers to never-married fertile women, spinsters beyond child-bearing age, and widows (VV 2022, pp. 1255-6). But all that this correlation shows is that the proportional fall in taxpayers between 1377 and 1381 was greater when the number of females who went missing from the tax returns was proportionally larger than the number of missing males. Since the females who were dropped included an unknown combination of spinsters beyond child-bearing age and widowed females as well as never-married fertile women, and furthermore the gender, age, and marital composition of the missing males and females may well have varied geographically, this correlation does not make the proportional fall in taxpayers a defensible proxy for the proportion of never-married women of child-bearing age, contrary to VV's claim.

4.2 Pastoralism and Female Age at First Marriage 1600-1837

Several issues arise in relation to VV's claim that evidence of a relationship between female age at first marriage (FAFM) and pastoralism in the period 1600-1837 supports their argument that the Black Death caused the EMP.

4.2.1 Relevance?

Our main criticism is that the existence of a relationship between FAFM and pastoralism in 1600-1837 can be explained in many ways that have nothing to do with the Black Death c. 1350. As we discuss in detail, a link between marriage behaviour and the regional distribution of pastoral agriculture in the 1600-1837 can be explained in terms of developments that occurred in the English agrarian economy in the 1600-1837 period, including the increasing regional specialization of agriculture, the growth of urban markets, and the adoption of new crops and techniques (Edwards and Ogilvie 2022, pp. 1237-8). There is thus no reason to believe that any such link had anything to do with the Black Death several centuries earlier. In their original article, VV provide no explanation of why, even if there were clear evidence that pastoralism had a causal effect on FAFM in a period 250-487 years after the Black Death (and, as we show, there is not), this evidence would support their argument that the EMP emerged in the aftermath of the Black Death. VV fail to address this objection in their reply. Thus VV simply ignore our main objection to their econometric analysis of the relationship between pastoralism and FAFM in 1600-1837, which is that it is not relevant to assessing any possible effect of the Black Death on fertility. Ignoring a criticism is not a rebuttal of it.

In fact, there is no causal relationship between FAFM and pastoralism in the period 1600-1837, as we show (Edwards and Ogilvie 2022, pp. 1238-42). VV do attempt to respond to our demonstration of this point, but the argument they advance cannot be sustained, as we shall now explain.

4.2.2 Inference with a Small Number of Clusters

One problem with VV's econometric analysis of the relationship between pastoralism and FAFM in the period 1600-1837 concerns the precision of the estimates. In the FAFM regressions that VV estimate, the number of clusters involved is so small that the wild cluster bootstrap (WCB) procedure recommended by Cameron et al. (2008) should be used to obtain the cluster-robust variance matrix rather than Stata's finite-sample adjustment, as we point out (Edwards and Ogilvie 2022, p. 1238). In our article we reported confidence intervals based on this procedure as well as confidence intervals based on Stata's finite-sample adjustment. VV imply that it was excusable for them not to have used the WCB procedure in their original article (VV 2013), on the grounds that the *boottest* command of Roodman et al. (2019) only became available after that article was published. But Cameron et al. (2008) was published, and the details of the WCB procedure were available, well before the publication of VV's original article in 2013. So VV could have used the WCB procedure in 2013. However, although the confidence intervals are wider when the WCB procedure is used than when they are based on Stata's finite-sample adjustment, Table 1 of Edwards and Ogilvie

(2022) and the associated discussion shows that the use of the WCB procedure does not lead to substantially different conclusions about the precision of the estimates from those obtained using Stata's finite-sample adjustment. The way in which the cluster-robust variance matrix is obtained is not the most important aspect of our criticism of VV's econometric analysis.

Although the appropriate way to make inferences when the number of clusters is small turns out to be relatively unimportant in this case, VV's discussion of the WCB procedure is remarkably confused. VV state that the confidence intervals obtained using the WCB procedure "rely on a numerical inversion of a bootstrap test that is known to be highly sensitive to the choice of grid search parameters" (VV 2022, p. 1257), and claim that Roodman et al. (2019) say such confidence intervals are problematic for this reason. This statement reveals that VV do not understand the issues involved. VV do not provide a specific reference for their claim that Roodman et al. say that confidence intervals are problematic in this case, and in fact Roodman et al. say no such thing. VV report *p* values of WCB tests that the point estimates of the pastoralism variables are zero, and appear to think that reporting *p* values is in some important way different from reporting confidence intervals (VV 2022, p. 1258). However, the 95 per cent confidence interval for a parameter β is constructed by finding all values of β^* such that the *p* value of the test of $\beta = \beta^*$ is greater than or equal to 0.05 (Roodman et al. 2019, p. 15). Reporting the *p* value of the test that $\beta = 0$ thus shows whether one particular value is or is not in the 95 per cent confidence interval. This, of course, provides no information about the precision of the estimate of β: such information requires the computation of a confidence interval in the manner described above. Roodman et al. (2019, p. 12) explain why standard errors are not good measures of precision when the WCB procedure is used. Provided that a large number of values of β^* are used in the search procedure, the confidence interval based on WCB tests can be robustly calculated.

VV go further when, on p. 1261 of their reply, they say

> [Edwards and Ogilvie] use the Roodman et al. (2019) Stata routine, which requires grid search parameters for the [confidence interval]. ... it is inexplicable why [Edwards and Ogilvie] would overwrite the default to specify a search with a lower bound of -1000. One cannot help but think of an attempt at 'fishing in the left tail' of the empirical distribution of the test statistic.

Our specification of a very large negative lower bound for the grid search parameter is not at all inexplicable. It is because we followed the appropriate procedure for the computation of confidence intervals which are based on the inversion of a hypothesis test. Roodman et al. (2019, p. 16) describe the procedure by which their Stata *boottest* command computes a confidence interval as follows:

> By default, *boottest* begins the search for confidence interval bounds by picking two trial values for [the parameter of interest], low and high, at which the *p*-value of the test < α [where the level of the test is 1-α]. *boottest* then calculates the p-value at 25 evenly spaced points between these extreme bounds. From these 25 results, it selects those adjacent pairs of points between which the *p*-value crosses α and then finds the crossover points via an iterative search. In [instrumental variable] applications .... when identification is weak, the confidence set constructed in this way may consist of more than one disjoint segment, and these segments may be unbounded ...

Roodman et al. (2019) therefore make it clear that in instrumental variable (IV) applications, which are the relevant ones for VV's econometric analysis, the confidence interval may comprise disjoint segments and be unbounded. Because of this possibility, use of the default grid search values sometimes results in *boottest* saying that a confidence interval could not be bounded and recommending that the default values should be overridden and replaced by values that cover a wider range. It is possible that even extremely large values for the range of the grid search fail to produce a bounded confidence interval. This is what happened in the cases where we reported unbounded WCB confidence intervals in Table 1 of Edwards and Ogilvie (2022): even enormous negative values for the lower bound of the grid search failed to yield a bounded confidence interval. We therefore truncated the bottom end of the search range at -1000 and reported the confidence interval as extending to -∞. There is nothing fishy or inexplicable about it. VV's failure to understand something that is clearly explained in

Roodman et al. (2019) does not permit them to accuse us of making a “dubious choice” (VV 2022, p. 1261). VV’s statement that “sensible [confidence intervals] cannot be disjoint” (VV 2022, p. 1261) further reveals that they do not understand the issues involved in inference with weak IVs.

We concur with VV that there is a case for two-way clustering of the regression errors by county and period, although we note that it is not clear how strong this case is when there is a very small number of clusters in one of the dimensions, as is the case here with only five periods (Cameron et al. 2011, MacKinnon et al. 2021). We also note that VV could have used two-way clustering in the regressions they reported in their original article, but did not. Does two-way rather than one-way clustering alter anything important in the results reported in Table 1 of our article? It typically results in much less precise estimates. For example, in equation (1.1) of Table 1 of Edwards and Ogilvie (2022) (which corresponds to equation (2) in Table 1 of VV 2022), the 95 per cent WCB confidence interval for *Pastoral 1290* with two-way clustering is [-27.33, -20.13] U [2.94, 8.36] compared to [1.12, 9.15] with one-way clustering. In equation (1.5) of Table 1 of our article (which corresponds to equation (3) in Table 1 of the VV reply), the 95 per cent WCB confidence interval for *Pastoral Marriage* with two-way clustering is [-20.09, -14.05] U [7.05, 12.21] compared to [6.70, 12.33] with one-way clustering. We do not, however, regard the changes in the precision of the IV estimates reported in Table 1 of our article to be a matter of any great consequence, because these estimates are obtained using IVs which are almost certainly invalid. We turn now to this point.

4.2.3 Are the IVs Valid?

In order for IV estimation to identify the causal effect of pastoralism on FAFM in the period 1600-1837, the IVs must influence FAFM only via their effect on the pastoralism regressor. In other words, valid IVs must satisfy the exclusion restriction. VV (2013) use ln(*daysgrass*), the logarithm of the number of days on which grass can grow in each county based on twentieth-century climate data, as the IV for pastoralism in the estimates they report in the main text. The number of days on which grass can grow is a measure of a county's suitability for grazing and thus should be positively correlated with its suitability for pastoral agriculture. However, although VV state what the exclusion restriction required for ln(*daysgrass*) to be a valid IV is (VV 2013, pp. 2253-4), they make no attempt to explain why ln(*daysgrass*) might satisfy it. The use of ln(*daysgrass*) as a valid IV for the pastoralism variables in the FAFM regressions is thus a matter of faith on VV's part. In their online Appendix, VV (2013) also use a second IV to obtain overidentified estimates. This second IV is crop suitability: the share of each county's area reaching a threshold yield for at least one of wheat, barley or rye, a measure of a county's suitability for arable agriculture which should therefore be negatively correlated with a county's suitability for pastoral farming. However, VV do not report the standard test of the overidentification restriction which their use of both ln(*daysgrass*) and crop suitability permits. When there is more than one IV, a minimum requirement for the IVs to be valid is that any overidentification restrictions should not be rejected.

VV's econometric analysis of the relationship between pastoralism and FAFM in the period 1600-1837 is flawed because the IVs they use are invalid. In Edwards and Ogilvie (2022, p. 1241) we provide reasons why it is likely that both ln(*daysgrass*) and crop suitability had a direct effect on FAFM that operated separately from any effect these IVs had via their effect on the pastoralism regressor, and hence that they are not valid IVs. We also report tests of the overidentification restriction, which show that at least one of ln(*daysgrass*) and crop suitability is an invalid IV.

VV (2022, pp. 1256-7) object to our use of crop suitability as an IV for the pastoralism variables in the regressions explaining FAFM in 1600-1837. The reason we used crop suitability as an IV as well as ln(*daysgrass*) is simple: VV themselves used both crop suitability and ln(*daysgrass*) as IVs in the Appendix to their original article. We did not "introduce" crop suitability, as VV claim: we simply used their own second IV. There is a very strong reason for using two IVs if they are available, which is that they permit the use of an overidentification test of the validity of the two IVs. The overidentification tests we report in Table 1 of our article show that at least one of crop suitability and ln(*daysgrass*) is an invalid IV. As we explain on p. 1241 of our article, it is unlikely that either crop suitability or ln(*daysgrass*) satisfies the exclusion restriction, so both IVs are probably invalid.

VV make much of our view that crop suitability is likely to be an invalid IV, saying that our use of it "[contradicts] their own discussion of possible endogeneity of this instrument" (VV 2022, p. 1256). VV fail to point out that we discuss the likely invalidity of crop suitability as an IV only after we have used both it and ln(*daysgrass*) to conduct overidentification tests which show that at least one of these IVs is invalid. They also make no mention of the arguments on p. 1241 of our article that neither ln(*daysgrass*) nor crop suitability are likely to satisfy the exclusion restriction. The most VV are able to do in response to these arguments is to inform us of their prior that ln(*daysgrass*) is more likely than crop suitability to satisfy the exclusion restriction (VV 2022, p. 1257), but they still fail to provide any argument in support of their belief. Our prior is different: we think that neither ln(*daysgrass*) nor crop suitability are likely to be valid IVs. In contrast to VV, we provide arguments to justify this belief: we expect a county's suitability for pastoral agriculture, which both ln(*daysgrass*) and crop suitability measure, to have a direct positive influence on FAFM in the 1600-1837 period because arable regions saw falling female labour productivity and stronger marriage incentives for women, while pastoral regions did not. VV (2022) ignore these arguments.

VV's view that crop suitability is not a valid IV for the pastoralism regressors was not mentioned when they used this IV themselves in the Appendix to their original article. Thus we are completely unpersuaded by their contention that the rejection of the overidentification restriction is because crop suitability is an invalid IV while ln(*daysgrass*) is a valid one. We see no reason to change our view that, because neither ln(*daysgrass*) nor crop suitability is likely to be a valid IV, it is not possible to obtain causal estimates of the effects of pastoralism on FAFM in the period 1600-1837. As we show in equations (1.4) and (1.8) of Table 1 in Edwards and Ogilvie (2022), the most that can be said on the basis of the data available is that there is a very small and poorly-determined positive association between pastoralism and FAFM in the period 1600-1837 (which, we reiterate, post-dates the Black Death by 250-487 years and thus cannot show anything about the relationship between pastoralism and FAFM in the aftermath of the Black Death).

4.2.4 Anderson-Rubin Tests?

In Table 1 of their reply, VV report *p* values of Anderson-Rubin (AR) tests of whether the coefficients of the pastoral variables are zero, based on two-way WCB variance estimators. These AR tests are robust to weak IVs, and, when a single IV is used for a single endogenous regressor, they are efficient provided that the IV is valid (Andrews et al. 2019). However, the null hypothesis of the AR test is that the coefficient of a particular potentially endogenous regressor takes a certain value *and* that the IVs are valid. It is therefore difficult to interpret the rejection of the null hypothesis by the AR test in Table 1 of VV (2022): it could be either because the coefficient estimate is different from zero or because the IV is invalid. As we have argued, ln(*daysgrass*), the IV used in Table 1 of VV (2022), is likely to be invalid, and use of the AR test is therefore likely to be misleading. In any case, the 95 per cent confidence intervals obtained using the WCB AR tests in Table 1 of VV are very similar to those we reported for equations (1.1) and (1.5) in Table 1 of our article, which correspond to equations (2) and (3) in VV's Table 1. Our disagreement with VV about the FAFM regression estimates has very little to do with the precision of the estimates. It is primarily about IV validity.

The results in Table 1 of VV's reply do not enable them to "resoundingly reject" (VV 2022, p. 1260) our criticisms of their FAFM regressions. VV ignore our argument that ln(*daysgrass*) is an invalid IV. Ignoring an argument does not amount to a rejection of it.

In Table 2 of their reply, VV also use the AR test to report *p* values for the point estimates of the pastoral variables in the two regressions we estimated using both crop suitability and ln(*daysgrass*) as IVs, equations (1.3) and (1.7) in Table 1 of our article. As we have already explained, we estimated these regressions because they permitted an overidentification test of the validity of the IVs. The overidentification restriction was rejected, and this is part of our argument that neither crop suitability nor ln(*daysgrass*) are likely to be valid IVs. The only conclusion that can be drawn from these two regressions is that they raise serious doubts about the validity of the IVs: certainly the estimates of the pastoral variables obtained from overidentified IV estimation when the overidentification restriction is rejected cannot be interpreted as giving estimates of causal effects. VV, however, interpret the very precise estimates of the effects of the pastoral variables obtained from their two-way WCB AR tests as showing that their original claims about a positive effect of past and contemporaneous pastoralism on FAFM 1600-1837 are robust (VV 2022,

p. 1260). This reveals a failure by VV to understand that we estimated these regressions in order to test the overidentification restriction, and once this restriction had been rejected, we attached no significance to the estimated effects of the pastoralism variables.

The greater precision of the AR tests for the two regressions in Table 2 of VV (2022) is, in fact, exactly what would be expected to result from the rejection of the overidentification restriction in these equations. Davidson and MacKinnon (2014) show that confidence intervals constructed from AR tests are typically too narrow when the overidentification restriction is rejected and thus suggest that a parameter is estimated more precisely than is actually the case. In equation (1.3) of Table 1 of our article, the two-way WCB 95 per cent confidence interval for the pastoral variable that is not robust to weak IVs is $[-\infty, -48.43] \cup [-0.18, 6.40]$, while in equation (1.7) it is $[-32.94, -19.3] \cup [2.59, 11.69]$. The corresponding 95 per cent confidence intervals obtained from the AR test are [3.13, 3.94] and [7.26, 13.17]. The narrower AR confidence intervals are a consequence of the rejection of the overidentification restriction, and thus reinforce our criticism that both crop suitability and ln(*daysgrass*) are not valid IVs. VV's belief that the AR tests support their econometric analysis because they suggest very precise estimates is incorrect. In fact, this feature of the AR tests supports our view that the IVs are not valid.

5. Conclusion

We identify fundamental flaws in the claim that the Black Death led to the emergence of the EMP, thereby freeing western Europe from Malthusian limits on growth. Careful examination, we show, reveals pervasive problems with that argument, affecting the basic premise, the underlying theoretical model, and the empirical analysis (Edwards and Ogilvie 2022).

The attempt by VV (2022) to re-establish their argument fails to address the multiple flaws we identified. They reiterate their original assertion that the EMP emerged in England after 1350, but fail to engage with the fact that historical demographers are widely divided on the issue, that there is no evidence that the EMP emerged in rural England in the aftermath of the Black Death, and that the data are too fragile for any definitive conclusions about the existence of the EMP in England before c. 1540. VV sidestep the logic of their own model, refusing to acknowledge the incontrovertible point that the factual inaccuracy of one key assumption – that women had to remain unmarried to work in pastoral agriculture – makes the model completely inapplicable to demographic behaviour in England after the Black Death. They make no attempt to justify the very special assumption about preferences which is the other key assumption of their model. They repeatedly refer to evidence on demographic behaviour centuries later than the Black Death with no attempt to explain how it could possibly be relevant to the aftermath of that pandemic. This weakness of their analysis also applies to much of the econometric evidence they adduce, but that evidence is, in addition, further vitiated by the inadequacy of their proxy for late medieval female celibacy and the invalidity of the IVs they use.

The test of an argument is whether it explains something that is believed to have taken place, advances a theoretical model whose assumptions fit the facts, deploys evidence pertinent to the matter at hand, and analyzes such evidence meticulously. VV's argument fails this test. There is no evidence that the Black Death led to the development of the EMP in England.

References

Allen, R. (1991). “The Two English Agricultural Revolutions: 1450-1850.” In B. Campbell and M. Overton, eds., *Land, Labour, and Livestock: Historical Studies in European Agricultural Productivity*. Manchester, Manchester University Press: 236-254.

Andrews, I., J. Stock and L. Sun (2019). “Weak Instruments in Instrumental Variables Regression: Theory and Practice.” *Annual Review of Economics* 11: 727-753.

Bailey, M. (1996). “Demographic Decline in Late-Medieval England: Some Thoughts on Recent Research.” *Economic History Review* 49(1): 1-19.

Bailey, M. (2007) *Medieval Suffolk: an Economic and Social History, 1200-1500*. Woodbridge, Boydell.

Bridbury, A. R. (1977). “Before the Black Death.” *Economic History Review* 30(3): 393-410.

Cameron, A. C., J. G. Gelbach and D. L. Miller (2008). “Bootstrap-Based Improvements for Inference with Clustered Errors.” *Review of Economics and Statistics* 90(3): 414-427.

Cameron, A. C., J. G. Gelbach and D. L. Miller (2011). “Robust Inference With Multiway Clustering.” *Journal of Business and Economic Statistics* 29(2): 238-249.

Campbell, B. M. S. (2000). *English Seigniorial Agriculture, 1250-1450*. Cambridge, Cambridge University Press.

Davidson, R. and J. MacKinnon (2014). “Confidence Sets Based on Inverting Anderson-Rubin Tests.” *The Econometrics Journal* 17(2): S39-S58.

Dennison, T. K., and S. Ogilvie (2014). “Does the European Marriage Pattern Explain Economic Growth?”, *The Journal of Economic History*, 74(3): 651-693.

Dennison, T. K., and S. Ogilvie (2016). “Institutions, Demography, and Economic Growth”, *The Journal of Economic History*, 76(1): 205-217.

Dyer, C. (1981). *Warwickshire Farming 1349-c. 1520: Preparations for Agricultural Revolution*. Dugdale Society Occasional Papers, 27.

Dyer, C. (2005). *Age of Transition? Economy and Society in England in the Later Middle Ages*. Oxford, Clarendon Press.

Edwards, J. and S. Ogilvie (2021). “Did the Black Death Cause Economic Development by ‘Inventing’ Fertility Restriction?” *Oxford Economic Papers* 74(4): 1228-1246.

Fenwick, C. C. (1998). *The Poll Taxes of 1377, 1379 and 1381: Part I*. Oxford, Oxford University Press.

Goldberg, P. J. P. (1986). “Marriage, Migration, Servanthood and Life-Cycle in Yorkshire Towns of the Late Middle Ages.” *Continuity and Change* 1(2): 141-169.

Goldberg, P. J. P. (1990). “Urban Identity and the Poll Taxes of 1377, 1379, and 1381.” *Economic History Review* 43(2): 194-216.

Hajnal, J. (1965). “European Marriage Patterns in Perspective.” In D. V. Glass and D. E. C. Eversley (eds.), *Population in History: Essays in Historical Demography*. London: Arnold: 101-143.

Hajnal, J. (1982). “Two Kinds of Preindustrial Household Formation System.” *Population and Development Review* 8(3): 449-494.

Hajnal, J. (1983). “Two Kinds of Pre-Industrial Household Formation System.” In R. Wall, J. Robin and P. Laslett (eds.), *Family Forms in Historic Europe.* Cambridge: Cambridge University Press: 65-104.

Hilton, R. H. (1975). *The English Peasantry in the Later Middle Ages*. Oxford, Clarendon Press.

Hilton, R. H. (1985). “Some Social and Economic Evidence in Late Medieval English Tax Returns.” In R. H. Hilton, *Class Conflict and the Crisis of Feudalism: Essays in Medieval Social History*. London, Hambledon Press: 253-267.

Humphries, J. and J. Weisdorf (2015). “The Wages of Women in England, 1260–1850.” *Journal of Economic History* 75(2): 405-447.

Kussmaul, A. (1981). *Servants in Husbandry in Early Modern England*. Cambridge, Cambridge University Press.

Laslett, P. (1988). “The Institution of Service.” *Local Population Studies* 40: 55-60.

MacKinnon, J., M. Nielsen and M. Webb (2021). “Wild Bootstrap and Asymptotic Inference with Multiway Clustering.” *Journal of Business and Economic Statistics* 39(2): 505-519.

Masschaele, J. (1997). *Peasants, Merchants and Markets: Inland Trade in Medieval England, 1150-1350*. London, Macmillan.

Mate, M. E. (1998). *Daughters, Wives and Widows after the Black Death: Women in Sussex, 1350-1535*. Woodbridge, Boydell.

Mate, M. E. (1999a). *Women in Medieval English Society*. Cambridge, Cambridge University Press.

Mate, M. E. (1999b). “Review of James Masschaele, Peasants, Merchants and Markets: Inland Trade in Medieval England.” *American Historical Review* 104(3): 975.

Poos, L. R. (1991). *A Rural Society after the Black Death: Essex 1350-1525*. Cambridge, Cambridge University Press.

Romer, P. (2015). “Mathiness in the Theory of Economic Growth.” *American Economic*

Review 105(5): 89-93.

Roodman, D., J. G. MacKinnon, M. Ø. Nielsen and M. D. Webb (2019). “Fast and Wild: Bootstrap Inference in Stata using Boottest.” *The Stata Journal* 19(1): 4-60.

Slavin, P. (2010). “Goose Management and Rearing in Late Medieval Eastern England, c.1250-1400.” *Agricultural History Review,* 58: 1-29.

Slavin, P. (2015). “Peasant Livestock Husbandry in Late Thirteenth-Century Suffolk: Economy, Environment, and Society.” In M. Kowaleski, J. Langdon, and P. R. Schofield, eds., *Peasants and Lords in the Medieval English Economy: Essays in Honour of Bruce M. S. Campbell*, Brepols, Turnhout.

Stone, D. (2003) “The Productivity and Management of Sheep in Late Medieval England”, *Agricultural History Review,* 51: 1-22.

Voigtländer, N. and H.-J. Voth (2013). “How the West ‘Invented’ Fertility Restriction.” *American Economic Review* 103(6): 2227-2264. Appendix at https://assets.aeaweb.org/asset-server/articles-attachments/aer/data/oct2013/20110925_app.pdf

Voigtländer, N. and H.-J. Voth (2022). “Reply to Edwards and Ogilvie: Did the Black Death Cause Economic Development by ‘Inventing’ Fertility Restriction?” *Oxford Economic Papers* 74(4): 1247-1263.

Watkins, A. (1989). “Cattle Grazing in the Forest of Arden in the Later Middle Ages.” *Agricultural History Review* 37(1): 12-25.

Whittle, J. (2017). “Servants in the Economy and Society of Rural Europe.” In J. Whittle, ed., *Servants in Rural Europe 1400-1900*. Woodbridge, Boydell and Brewer: 1-18.