DISCUSSION PAPER

April 2007 ■ RFF DP 07-23

What Drives Long-Term Biodiversity Change?

New Insights from Combining Economics, Paleo-Ecology, and Environmental History

Nick Hanley, Konstantinos Angelopoulos, Dugald Tinch, Althea Davies, Fiona Watson, and Edward B. Barbier

The Frontiers of Environmental Economics

Hosted by Resources for the Future
1616 P Street, NW
Washington, DC 20036
202.328.5000
www.rff.org/frontiersconference

The Frontiers of Environmental Economics Conference, Washington DC, February 2007

What Drives Long-Term Biodiversity Change? New Insights from Combining Economics, Paleo-Ecology, and Environmental History

Nick Hanley, Konstantinos Angelopoulos and Dugald Tinch,
Economics Dept., University of Stirling, Stirling, Scotland.

Althea Davies,
AHRC Centre for Environmental History, University of Stirling, Scotland

Fiona Watson,
History Department, University of Dundee, Scotland.

Edward B. Barbier, Dept. of Economics and Finance, University of Wyoming.

Address for correspondence: Nick Hanley, Economics Department, University of Stirling, Stirling FK9 4LA, Scotland, UK. Email n.d.hanley@stir.ac.uk.

We thank the Leverhulme Trust for funding the research on which this paper is based under grant number F/00241/C. We also thank participants at the First Conference on Early Economic Developments, Copenhagen, 2006 for helpful comments, especially Paul Seabright and Gregory Clark. Thanks to Alistair Hamilton and Alasdair Ross for research assistance and to Elias Tzavalis, Chris Smout and Brian Huntley for comments.

Abstract

This paper presents an econometric analysis of the relationship between economic and environmental factors and biodiversity change in Scotland over the period 1600-2000. Our main hypothesis is that economic development, as captured by changes in prices, technological improvements and management intensity, is a significant determinant of long-run biodiversity change. The measure of biodiversity used here is an estimate of plant species diversity, constructed by the authors using paleo-ecological analysis of pollen remains. We assemble a new data set of historical land use and price data over 11 sites during this 400 year period; this data set also includes information on long-run climate change and extreme weather events, as well as changes in agricultural technology, land tenure and land ownership. A panel model is then estimated, which controls for both supply and demand shifts over time. Our main result is that agricultural prices, which determine livestock numbers, do indeed impact on biodiversity, with higher prices leading to lower diversity scores, due to their influence on production. No significant direct effects of either changes in technology or climate are detected.

Keywords: agriculture, biodiversity, climate change, paleo-ecology, panel models, instrumental variables.

JEL codes: C33, N53, O13, Q57

1. Introduction

The state of a nation's "biodiversity" has emerged as an increasingly important indicator of environmental health (Wilson, 1988). Biodiversity incorporates the range and abundance of plant and animal species, the interactions between them, and the natural systems that support them (Armsworth et al, 2004). Whilst many measures of biodiversity exist, the number of different species existing in a given area is an important component of most indicators, and this is the concept used in this paper. Biodiversity can be expected to change over time as ecosystems evolve, partly in response to exogenous shocks. What interests us in this paper is quantifying the long-run relationship between biodiversity and the functioning of the economic system: in particular, we focus on agricultural change as a driver of biodiversity change.

Threats to biodiversity from human activity are usually thought of by biologists in terms of habitat loss, degradation, and fragmentation, harvesting and human-induced climate change (Orians and Soule, 2001). Addressing these threats at both the theoretical and empirical level has been an important theme in environmental economics work in the recent past, as evidenced for instance in work on drivers of rainforest loss (Barbier and Burgess, 2001). But at the empirical level, this work has been limited to looking either at rather recent cross-sectional data (eg species loss by country) or at rather short-duration time series data, typically looking no further back than the 1970s.

The main contribution of this paper is the assembling and analysis of a unique data set which allows econometric modeling of one measure of biodiversity change (plant species numbers) as a function of economic development in an agricultural economy over a much longer, 400 year period. This data set is assembled using inputs from economic history and palaeo-ecology for a sample of upland sites in Scotland. We estimate a structural model which is based on the dominant ecological theory about what drives plant species change in the uplands of North-West Europe, namely changes in grazing pressure from livestock (Anderson and Yalden, 1981; Thompson et al, 1995; Palmer, 1997; Fuller and Gough, 1999[1]). Given the lack of historical data on livestock numbers, we illustrate how livestock prices may be used instead of grazing pressure as a determinant of long-run biodiversity impacts. We hope that these results can help explain past changes in biodiversity, and also help cast some light on how the uplands can be managed in the future to achieve national biodiversity targets. In what follows, we first describe the approach taken to assembling the data set, before setting out some modeling considerations. Results are then presented and discussed.

2. Approach to data collection

Virtually all of the data used in the model had to be obtained from both documentary primary sources and palaeo-ecological analysis by the research team. The first requirement was to select the sites to be used for data collection. Sites were intended to represent a range of biogeographical zones in the Scottish uplands. This was an iterative process, involving identifying sites with historical potential (ie sites where there was a reasonable chance of obtaining enough documentary sources), alongside fieldwork to seek suitable peat deposits (to obtain intact, undisturbed historic pollen sequences), and then final joint site selection. Sites are shown in Figure One. All sites were predominantly upland livestock farms, with limited arable cropping.

The second need was to construct a time series for a biodiversity index for each of our sites. This was accomplished by focusing on a proxy for plant diversity using palaeo-ecological techniques (Birks and Line, 1992). We refer to this measure below as B_{it}, the estimated plant species count at site i in time period t. This involved taking pollen samples from peat cores, dating these using a combination of radiocarbon (^{14}C) and lead-210 techniques, and identifying and quantifying the plant types present in the peat sequence, thus effectively reconstructing

[1] For example, Fuller and Gough (1999) argue that increases in sheep numbers in upland Wales from 1970s to 1990s have almost certainly caused reduction in habitat quality for ground-nesting birds such as grouse and waders, partly through the effects on plant cover, leading to a decline in bird numbers.

vegetation change through time[2]. The pollen signal is sensitive to and records vegetation cover within a radius of up to 1 kilometer at our sites. The pollen analysis also allows us to see how the composition of species changed through time at individual sites. Figure 2 shows example pollen data from four sites[3]. As can be seen, the species count does not change linearly or monotonically over the time period.

The third need was to construct a historical and cross-sectional data base of agricultural land management. Cattle and sheep grazing was the dominant agricultural land use in the sites we investigated in rural Scotland over the period in question, but we expect impacts on biodiversity to depend on how intensively land was managed - particularly on stocking patterns - and what technology was available and utilized (eg new breeds of sheep which exert different grazing pressure than older breeds). A contemporary study of agricultural impacts on upland plant diversity would focus on grazing pressure, measured by livestock units per hectare (ha). Unfortunately, the records of livestock numbers and area grazed on individual farms are very patchy, and official data was only collected on this from the 1860s onwards, and then only at a higher level of spatial aggregation. Individual estate records typically do not record either the area being grazed or the total number of livestock at individual sites. We thus cannot use a traditional grazing intensity measure. Instead, we reconstruct a time series of prices for livestock and crops by region, and we show that higher prices of livestock (for meat) and other products (e.g. wool), ceteris paribus, would motivate farmers to increase their herds as a normal supply response (see the Appendix). However, we are able to represent technological change, by creating count variables for recorded instances of new breeds or new agricultural techniques such as liming or

[2] Note that this is an estimate of the number of plant species since not all plant species can be distinguished from their pollen remains, whilst the dating of each sample is also an estimate. Samples were taken from small flushes and mires, rather than large blanket peat or raised mire sites.

[3] Although we only run the model from 1600 to 2000 due to the lack of historical sources prior to 1600, the pollen data is in fact available back as far as 5500 years ago for one of our sites.

the growing of fodder crops being introduced at each of our sites[4]. Distinct changes in farm management such as enclosure are also recorded[5], whilst we are also able to record the degree of utilization of each through a typical farming year, from abandonment to year-round cultivation . We might also anticipate that changes in land ownership, size of holding (due to farm amalgamations, which occurred at most of our sites) or tenancy might cause (un-observed) changes in land use, thus dummy variables were also constructed to measure such changes.

The historical data was collected, firstly, from the estate papers (archives of material relating to the landholdings of particular noble families and encompassing a wide range of material generated by, most usually, the owners, their estate officials and lawyers, and, less frequently, their tenant farmers) relating to each site, ie. Scott of Buccleuch (Bush of Ewes and Foulshiels study sites), Campbells of Glenorchy/Breadalbane (Leadour and Corries), Sutherland (Glenleraig, Rogart), Grant of Freuchy (for Abernethy site at Rynuie). These are mostly to be found in the National Archives of Scotland in Edinburgh, although some of the Sutherland papers are still held *in situ*. As well as searching for evidence specifically for each site and of the kind of quantitative and qualitative detail necessary for the model (eg. animal breeds, prices, ownership patterns, changes in land management), we also collected more general material, both spatially (ie. covering neighbouring farms to act as a comparison, corroboration, or fill data gaps) and socially (ie. material of a more general nature to illuminate the wider estate/regional context

[4] The 18th century saw the gradual replacement of native sheep breeds in Scotland with two new "imports": the Cheviot and Blackface sheep. These rapidly spread through Scotland during the 18th and 19th centuries (Carlyle, 1979). Cheviots were favoured for the higher price their wool could command, whilst blackfaced sheep were hardier than native breeds and could be over-wintered on the hill. Both new breeds also had bigger carcase weights than native breeds. Due to differences in their grazing behaviour, dry matter intake and length of time on the hill, the introduction of both breeds could be expected to have an effect on plant cover. Cheviots reached their peak in terms of geographic coverage of Scotland in the 1860s – 1870s, from when they were gradually replaced by blackfaced sheep and cross-breeds. Their decline is attributed to an over-extension of geographic range, falling wool prices due to imports from Australia and New Zealand, and changing preferences for sheepmeat. Note that we were unable to model the effects of wool prices due to a lack of data.

[5] Enclosure is only noted at three of our sites: Abernethy in 1763, Rogart between 1781-1800 and Corries in 1841. Enclosure has been argued to have been responsible for a major increase in Scottish agricultural productivity (eg Devine, 1994) but this has been questioned by others, who pinpoint the late 17th and early 19th centuries as being more associated with major increases in output, with output stagnating or even falling during the main period of enclosure (eg see Douglas, 2004; Whyte, 1979; and, for a similar viewpoint in an English context, Allen (1999)).

within which our sites fitted). Finally, because of the lack of price data at a site-specific level, neighbouring estate papers were consulted in addition to provide a credible dataset at the regional level. A timeline of significant external events (eg. national or significant regional famine, major periods of warfare) was also constructed from secondary sources to act as a wider context for site specific activity.

Prices for livestock (sheep and cattle) were taken from estate papers, as noted above, for the early period, and from secondary sources for later periods. In the former case, these prices often relate to local livestock auctions, and price series were assembled for each region in our analysis. Despite these regional price series being rather patchy and incomplete, they show that regional prices closely tracked each other over the period 1580-1880. After 1880, we assume a single national price exists for sheep and for cattle. Prices for arable crops (which although a minor part of the farms being studied nonetheless provide additional income and direct subsistence) were taken from the "fiars" prices available from 1626-1780 in Gibson and Smout (1995a). Later figures for these prices were found in the General Records Office. Crops focussed on were those of most relevance to our case study sites, namely bere (an early type of barley). Fiars prices were "declared" by regional sheriff courts each year as a "..just assessment of the prevailing winter price for each type of grain grown and traded". Their use was as officially-sanctioned prices in settling a range of contracts and bargains. According to Gibson and Smout (1995b), .." *comparisons with known transactions prices tend to confirm their reliability*". Whilst regional series exist for these prices, Gibson and Smout (1995b) argue that regional grain markets were well-integrated by the late seventeenth century.

Finally, information was needed on environmental factors likely to influence biodiversity change. Since no long-term time series on climate is available at even the national level for Scotland, we use English data for precipitation and temperature. However, an "extreme weather events" dummy variable is also constructed for each site, to represent weather events such as floods or droughts that were unusual enough to be recorded in our historical documents (although in the early period this record is rather fragmentary).

Table 1 summarises the data series available for use in the model.

3. Modeling strategy

Adequate data are not available to conduct either a time-series analysis of drivers of biodiversity change at one site, or a cross-sectional analysis across all sites in a given time period[6]. Instead, we use panel data techniques to allow variability across time to be considered jointly with variability across space. The model we are interested in estimating can be written as:

$$B_{it} = \alpha B_{i,t-1} + bQ_{it} + S_{it}\delta + c_i + u_{it} \qquad (1),$$

where B_{it} is our measure of biodiversity, Q_{it} is a measure of the numbers of livestock which farmers keep and thus the preferred indicator of grazing pressure, S_{it} includes other observed variables that are also thought to affect biodiversity, c_i are site-specific (fixed) effects relating to biodiversity levels (such as soil type and elevation), u_{it} is the idiosyncratic error term and α, b, δ are parameters to be estimated.

Our estimate of biodiversity is state dependent. Past vegetation composition and land-use influence current ecology, but the rates at which plants respond to change may differ between species. The ecological argument is thus in favour of including the past species count as a determinant of the current number of species at a site. We therefore include a lagged term for species number, $B_{i,t-1}$ as a predictor of B_{it}. We expect higher values for $B_{i,t-1}$ to result in higher values of B_{it}. However, our main interest here lies on the effect of economic variables on

[6] This is symptomatic of the paucity of records existing for many aspects of land management in Scotland over the period.

biodiversity and primarily on the effect of the variable Q_{it} on biodiversity. As we note below, however, we cannot directly observe Q_{it}. We would expect, as noted in Section 1, that higher livestock densities are in general associated with lower levels of plant diversity, although we also allow for other influences. For instance, we include in S_{it} management variables, such as *sizechange*, *mgtchange* and *mgtinten* : the first of these represents whether farm amalgamations occurred in a time period, and we know from historical analysis that such amalgamations are sometimes linked to changes in management. *Mgtchange* is a count variable which represents changes such as enclosure and large-scale draining. *Mgtinten* represents how much of the year the site was actively managed for agriculture, from abandonment to year-round use. We also include in S_{it} some historical, technological and climatological variables that may also affect biodiversity. These are *andisease*, *annewbread*, *extrweather* and *extrcivil*. These represent incidents of animal disease (associated with falling stocking densities), the introduction of new breeds (which might, per head, have higher grazing demands), extreme weather events that were sufficiently unusual to be recorded, and extreme civil events such as civil war, which might disrupt supply chains and take labour away from farms. Finally, site fixed effects are included to represent the importance of factors such as soil type and altitude on biodiversity change.

To allow for the likely non-alignment in time of diversity and historical information, and to handle the relative paucity of historical information on land use change, we decided to construct 20-year "time slices" over the 400-year study period. The model thus analyses change from one twenty-year period to the next. Where multiple responses are available on a variable within a twenty-year period, we simply construct a mean score (for quantitative independent variables) or a count (for discrete independent variables). However, the paucity of historical sources available means we often encounter gaps in even this 20-year averaged data for some variables. Although the aim of pollen analysis was to provide a sample every 20 years, more "observations" are available in recent, near-surface sediments due to the relative lack of compaction and decay in upper peat compared with older, deeper sediments. As this varies between sites, our final dataset is not balanced and we can finally use (taking into account lagged-variables requirements) 2 to 17 observations from the each of the 11 sites, yielding a total of 119 observations.

Our modeling strategy is as follows. First, we control for the site-specific effects directly by including a dummy for each site. We then turn to the variable Q_{it}. As noted in the previous section, we cannot observe the number of animals on each of the sites in each time period; this historical information simply does not exist. As shown in the appendix to this paper, however, overall herd size per ha, and thus grazing pressure, will increase with the market price of livestock (e.g. for meat), $dQ_{it}/dp_{it} > 0$. This suggests that we can use instead of Q_{it} in equation (1) either the price of cattle (denoted as *pcatttle* in Table 1) or the price of sheep (denoted as *psheep*). As discussed in the previous section, historical regional series exist for these livestock prices, but the regional prices closely track each other over 1580-1880 and after 1880 there is essentially a single national price for sheep and cattle. This implies that the observed market prices available for our analysis is effectively a single prevailing price for cattle and sheep. In other words, these prices change through time but are taken to be the same for all sites, i.e. $p_{it} \approx P_t$ over the period of our analysis. The most likely reason why all sites face the same livestock price is that regional markets are well-integrated and prices are therefore endogenously determined by supply and demand. Nonetheless, we expect each individual farmer to treat the market price as exogenous, and to act accordingly. As shown in the appendix, in response to a rise in the price of livestock (for meat) the farmer will want to sell more livestock, and thus will increase the existing herd size on the farm. The result of this supply response, according to (1), should ceteris paribus be a fall in B_{it}. But unless we correct for the endogeneity of the observed *pcattle* and *psheep*, this effect is uncertain in our analysis.

In sum, the most important econometric issue with substituting our observed livestock prices *pcattle* and *psheep* for Q_{it} in (1) is that prices are endogenous in this regression, so that their effect is not immediately identified as a demand or a supply effect. If we could assume the existence of a supply equation,

$$Q_{it} = \eta P_t + S_{it}\theta + e_{it} \qquad (2)$$

then an increase in prices would result in an increase in the number of animals per ha and hence a decrease in the number of species (a fall in B_{it}). However, the equilibrium prices that we observe historically are most likely an endogenous outcome, determined jointly with quantity, so that the effect of prices in (2), and hence in (1) may be affected by reverse causality, and therefore is not identified. To make this clear, substitute (2) in (1) to get (note that the supply shifters in (2) are essentially the variables already included as S_{it} in (1)):

$$\begin{aligned} &B_{it} = \alpha B_{i,t-1} + (b\eta) P_t + S_{it}(\delta + b\theta) + c_i + u_{it} + be_{it}, or \\ &B_{it} = \alpha B_{i,t-1} + \beta P_t + S_{it}\gamma + c_i + v_{it} \\ &where \\ &\beta = b\eta, \gamma = \delta + b\theta, v_{it} = u_{it} + be_{it} \end{aligned} \qquad (3)$$

Therefore, α, β, γ are the parameters we can estimate in (3). Since P_t is correlated with e_{it} in (2), because of simultaneity, P_t will also be correlated with v_{it} in (3).

Our approach to identify the effect of P_t in (3) is essentially the method used to identify P_t in a supply equation like (2). That is, we use demand shifters that are correlated with prices, but uncorrelated with e_{it} (and hence with v_{it}), as instruments in IV methods to estimate (3). In this way, since P_t is identified in (2), we expect η to be positive and thus a negative β will imply a negative b. As demand shifters we use the variables: *pbere* (as the price of a substitute in consumption: none of our sites engaged in significant grain production, due to their locations), *garrison* (as the installing of military garrisons during periods of war and civil unrest may have been a new source of demand), *union* (as the relaxing of trade barriers between England and Scotland following the Act of Union in 1707 may have provided increased demand from a new market), *popenglish* (for increased demand from consumers in England), and *refrigeration* (as the advent of refrigerated transport in the 1890s meant that consumers could substitute imported meat from the New World for Scottish meat). These variables can be thought of as unrelated with either e_{it} or u_{it}, conditioning on the right hand side variables in equations (2) and (3) and can thus be used together with the variables in S_{it} as instruments for the prices. In any case, the validity of the instruments will be tested by over-identification tests.

The final issue we deal with is the presence of the lagged endogenous variable as a regressor. This implies that (3) will not satisfy the strict exogeneity assumption needed for the fixed effects estimator to be consistent, as v_{it} will be correlated with future realizations of $B_{i,t-1}$. In such dynamic models, the usual approach is to exploit *sequential* moment restrictions, i.e. the fact that the error term is correlated with leads but not with lags of $B_{i,t-1}$, and use the latter as instruments in IV methods. As the main interest here lies in consistently estimating (primarily) β and (also) γ, we deal with potential biases introduced by $B_{i,t-1}$, by using $B_{i,t-2}$, along with the demand shifters and the variables in S_{it} to instrument $B_{i,t-1}$ (see e.g. Wooldridge, 2002, chapter 11, for panel data models without the strict exogeneity assumption).

4. Results

The results are presented in Table 2. The first two columns present results using *pcattle* for P_t in (3), and the following two columns we use *psheep* for P_t. As the two prices are highly correlated (the correlation coefficient is 0.99) it makes little sense to include them together in the regression. The variables $B_{i,t-1}$, *pcattle* and *psheep* are treated as endogenous and the excluded instruments in these regressions are $B_{i,t-2}$, *popenglish*, *war*, *union*, *refrigeration* and *pbere*. Columns (1) and (3) presents 2SLS results while columns (2) and (4) report results obtained by Fuller's (1977) modified LIML, with a = 1, as it has been found in simulation studies to be more robust to potentially weak instruments (the potential biases due to weak instruments are much smaller with LIML, see e.g. Andrews and Stock, 2005, and Stock and Yogo, 2005).

Before discussing the results, we note that the model does well with respect to the diagnostics for the validity and relevance of the instruments. In particular, we first see that the Sargan over-identifying tests clearly support the null that the instruments are uncorrelated with the structural error term. In addition, the Anderson (1984) canonical correlations and the Cragg and Donald (1993) tests reject the null of under-identification. To further examine instrument relevance, we report the first stage F-statistics (of the test that the joint effect of the excluded instruments on the endogenous variable is zero in the first stage regression) and Shea's (1997) first stage "*partial R-squared*". Both present strong evidence of high correlation of the instruments with the endogenous variables (especially with prices). The Stock and Yogo (2005) tests for weak instruments suggest that the first stage correlations may introduce biases in the 2SLS regressions but not in LIML regressions and thus favour Fuller's LIML estimator. In any case, we do not find important differences between the 2SLS and LIML estimates. Finally, Shapiro and Wilk (1965) tests for normality suggest that the residuals from the structural equations are in all regressions normally distributed.

The results show that higher prices for both sheep and cattle imply lower levels of biodiversity over time and across sites. The implication is that the rise in the price in livestock markets for “meat on the hoof” means that the farmer will want to sell more livestock, and thus will want to increase the existing herd size. This response seems to “confirm” modern ecological thinking about the likely effects of overgrazing on fragile upland ecosystems. It is interesting to note the implication that increased sheep grazing (as captured by increases in the price of sheep) has been much worse for biodiversity than increased cattle grazing (as captured by increases in the price of cattle). However, we have not been able to identify thresholds up to which higher grazing pressures actually increase diversity (we have tried to include squared prices in our regressions as well, but these have turned out not to be significant). The only other variable that emerges as significant is the intensity with which sites area managed year-round; results show that abandonment of sites reduces biodiversity. Neither technological innovations nor extreme weather events seem to matter to our measure of biodiversity. Finally, in accord with expectations, it can be seen that higher species numbers in preceding periods is associated with higher species numbers in subsequent periods – there is a biological inheritance effect present in the data.

Since most of the variables in S_{it} are not significant, we repeat the regressions in columns (2) and (4) by keeping only *annewbreed* and *mgtinten* to check whether the estimates for the main variables of interest are affected by the inclusion of irrelevant variables (the former variable was retained since there has been considerable interest in the effects of new breeds on biodiversity). The new results are reported in columns (5) and (6). As may be seen, this produced no major changes to the results noted above.

[8] This implicit price will include, among other things, any loss of productivity of grazing land due to adding more livestock, the “interest” costs implied by “waiting” for young animals to achieve ideal weights and size for meat production, or wool in the case of some sheep, and the opportunity cost of maintaining land for grazing purposes over the long term. In this sense, (A.1) captures some of the dynamic aspects of a standard inter-temporal renewable resource stock management problem. This is seen more clearly by noting that the definition of current net additions to the herd is $n_{it} = g\left(Q_{t-1}\right) - h_{it}$, where the first term represents the biological reproduction of the surviving herd and h_{it} is any “offtake” of animals by the farm for its own consumption. Thus c_{it} corresponds closely to the shadow price of the current stock of animals

5. Conclusions

This paper has set out to investigate the effects of economic development, as mirrored in agricultural land use, on a measure of biodiversity over a 400-year period. We constructed a panel of estimates of plant diversity across space and time using pollen analysis, and assembled a data set of prices, land use change, technological improvements and changes in property rights. Panel regression analysis was then used to explore relationships between the diversity estimate and these economic and social drivers. The main conclusions that emerged were that agricultural prices exerted significant influences on diversity over the period 1600-2000, as did the extent to which sites were farmed year-round. However, no significant effects were found for climatic variables, or for extreme civil events or tenant change. Our results might thus be seen as confirming the ecological idea that rising grazing pressures is bad for biodiversity; but also suggests that land abandonment reduces diversity too. Both of these findings show that over the long run, human-induced biodiversity change was significant for these sites.

However, perhaps the approach and process behind this research is more interesting than the results. We know of no other similar combination of historical, palaeo-ecological and economic analysis to look at this or similar issues. Despite considerable gaps in the data (due in part to the paucity of historical records in Scotland for the early period), we were able to test whether change in biodiversity has been unidirectional over time, and what effects economic, social and environmental factors had on this. Problems of course exist. The first is simply that of missing information, most importantly perhaps on the number of animals grazed on our sites over time. We also note the problems in transforming historical information into a form suitable for quantitative analysis; for example, in terms of changes in farm management. Much detail is lost from the historian in transferring this information into a quantitative form useable in a regression model. To the ecologist, our measure of "biodiversity" would cause problems, in that it treats all observable species as equal (rather than placing a higher weight on, for instance, native or representative species relative to introduced species), whilst the pollen record cannot perfectly distinguish between species. Requiring a matching of historical and palaeo-ecological information has also caused difficulties. Where the historical data is relatively rich (eg the 17th century) the pollen record is rather poor, militating against a time series analysis for each site. In other periods, it is the lack of historical data that frustrates the analyst: historians are well-used to dealing with such gaps, but economists typically look for "full and complete" datasets before embarking on econometric research. This requirement would have stymied inter-disciplinary work of this kind if rigorously enforced.

Appendix

At each upland estate site investigated in this paper, the principal agriculture farm use has been rearing livestock (sheep and/or cattle), with only limited arable cropping (oats and/or barley). The main long-run impact on biodiversity at these sites has therefore occurred through increased intensity of livestock use, or grazing pressure, on estate farm land (see equation (1)). However, the lack of historical data for upland farms in Scotland to compute the main explanatory variable of livestock numbers per hectare, Q_t, prevents estimation of equation (1). The following appendix employs a model of farm-level behavior to illustrate how livestock prices may be used instead of grazing pressure as an explanatory variable in the long-run biodiversity relationship (1).

All farms are privately owned and managed, and it is assumed that the objective of the private owner at each site, in any time period *t*, is to maximize rent per ha through choice of land use for grazing sheep and/or cattle. Sheep and cows are raised for meat, and both livestock are sold in local markets as "meat on the hoof". Certain sheep breeds may also produce wool, which is also sold. Since crop production is incidental to each farm, crop land use decisions will be ignored, with little effect on the subsequent analysis. For the farm owner at each site *i*, the rent objective function can be specified as

$$\max \pi = \pi\left(p_{it}, c_{it}, w_{it}\right) = \max_{n_{it}, X_{it}} p_{it} f\left(Q_{it-1} + n_{it}, X_{it}\right) - c_{it} n_{it} - w_{it} X_{it}, \qquad \text{(A.1)}$$

where *f* is the per hectare (ha) production function for livestock output (e.g. meat, and in the case of sheep, wool) and has the normal concave properties. The price of this output sold in local markets is p_{it}. Production of this marketable output will depend on the (surviving) livestock herd from the previous period, $Q_{it\text{-}1}$, plus any net additional animals added to the herd in the current period, n_{it}. However, the farm will also sell some of its livestock, Q_{it}^{s}, as "meat on the hoof" in local markets. We assume that such net increases in the farm's current herd size per ha is from its own stock (e.g. through breeding), and thus c_{it} represents the implicit price or opportunity cost of doing so.[8] Finally, X_{it} is a vector of other inputs (e.g. labor, feed) used in raising livestock, with a corresponding vector of input prices w_{it}.

Thus the farm's current stocking rate per ha will be $Q_{it} = Q_{it-1} + n_{it}$. Since the surviving livestock herd, Q_{it-1}, is predetermined, the increased in current grazing pressure will be determined by the demand for net additions to the herd, n_{it}. The optimal demand function $n_{it}\left(p_{it}, c_{it}, w_{it}\right)$ must satisfy the necessary and first and second-order conditions for maximizing (A.1), which suppressing all other arguments except p_{it} and c_{it} are

$$p_{it} f'\left(n_{it}\left(p_{it}, c_{it}\right)\right) - c_{it} \equiv 0,\ p_{it} f''\left(n_{it}\left(p_{it}, c_{it}\right)\right) \leq 0 . \qquad \text{(A.2)}$$

These conditions are an identity in p_{it} and c_{it} since $n_{it}\left(p_{it}, c_{it}, w_{it}\right)$ must satisfy the necessary conditions for profit maximizing for all values of these prices. Assuming that the maximum is regular so that the second-order condition is not zero, then differentiating the first-order condition with respect to p_{it} and rearranging yields

$$\frac{dn_{it}\left(p_{it}, c_{it}\right)}{dp_{it}} \equiv -\frac{f'}{p_{it} f''} > 0 . \qquad \text{(A.3)}$$

The farm owner will increase the herd in the current period if the price of livestock (for meat) and their products (e.g. wool) rise in local markets. It follows that overall herd size per ha, and thus grazing pressure, will also increase with market price, $dQ_{it}/dp_{it} > 0$, and from equation (1), there will be a corresponding decrease in the number of species, $dB_{it}/dp_{it} < 0$.

In essence, this behavior is a supply response of the farmer. The rise in the price of cattle and sheep in livestock markets for "meat on the hoof" means that the farmer will want to sell more livestock, and thus will want to increase the existing herd size.

References

Allen R.C. (1999) "Tracking the agricultural revolution in England" *Economic History Review*, LII, 2 (209-235.

Anderson, T.W. (1984): *Introduction to Multivariate Statistical Analysis*, 2d edition, New York: John Wiley & Sons.

Anderson P. and Yalden D.W. (1981) "Increased sheep numbers and the loss of heather moorland in the Peak District, England" *Biological Conservation*, 20, 195-213.

Andrews D. W. K. and J. H. Stock (2005): Inference with weak instruments, Cowles Foundation Discussion Paper No. 1530.

Armsworth P.R., Kendall B.E. and Davis F.W. (2004) "An introduction to biodiversity concepts for environmental economists" *Resource and Energy Economics*, 26, 115-136.

Barbier E.B. and Burgess J.C. (2001) "The economics of tropical deforestation" in Hanley N and Roberts C (eds.) *Issues in Environmental Economics*. Oxford: Blackwells.

Birks, H.J.B. & Line, J.M., 1992. 'The use of rarefaction analysis for estimating palynological richness from Quaternary pollen-analytical data', *The Holocene*, 2, 1-10.

Carlyle, W.J. (1979) "The changing distribution of breeds of sheep in Scotland, 1795-1965" *The Agricultural History Review*, 27 (I), 19-29.

Cragg, J.G. and S.G. Donald (1993): Testing Identfiability and Specification in Instrumental Variables Models, *Econometric Theory*, 9, 222-240.

Devine T (1994) *The Transformation of Rural Scotland: Social Change and the Agrarian Economy, 1600-1815*. Edinburgh: Edinburgh University Press.

Douglas C. (2004) "Enclosure and agricultural development in Scotland" Paper to the Economic History Society conference.

Fuller W. A. (1977): Some properties of a modification of the Limited Information Estimator, *Econometrica*, 45, 939-954.

Fuller R.J. and S. J. Gough (1999) "Changes in sheep numbers in Britain: implications for bird populations". *Biological Conservation* 91 (1) 73-89.

Gibson A. J. and Smout T.C. (1995a) *Prices, food and wages in Scotland, 1550-1780*. Cambridge: Cambridge University Press.

Gibson A. J. and Smout T.C. (1995b) "Regional prices and market regions: the evolution of the early modern Scottish grain market" *Economic History Review*, XLVIII, 258-282.

Orians GH and Soule ME (2001) "Introduction" in Orians and Soule (eds) *Conservation Biology: Priorities for the next decade":* Washington DC: Island Press.

Palmer, S.C.F. (1997). Prediction of the shoot production of heather under grazing in the uplands of Great Britain. *Grass and Forage Science* **52**, 408-424.

Shapiro, S. S. and M. B. Wilk (1965): An analysis of variance test for normality, *Biometrika*, 52, 591-641.

Shea, J. (1997): Instrument relevance in multivariate linear models: A simple measure, *Review of Economics and Statistics*, 49, 348-352.

Stock, J. H. and M. Yogo (2005): Testing for Weak Instruments in Linear IV Regression, Ch.5 in J.H. Stock and D.W.K. Andrews (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Cambridge University Press.

Thompson D.B.A., MacDonald A.J. and Galbraith C.A. (1995) "Upland heather moorland in Great Britain: a review of international importance, vegetation change and some objectives for nature conservation" *Biological Conservation*, 71, 163-178.

Whyte I (1979) *Agriculture and Society in Seventeenth Century Scotland.* Edinburgh: John Donald Publishers Ltd.

Wilson E.O. (1988) *Biodiversity*. Washington DC: National Academy Press.

Wooldridge J. (2002) *Econometric analysis of cross section and panel data"* . Cambridge., Ma. :MIT Press.

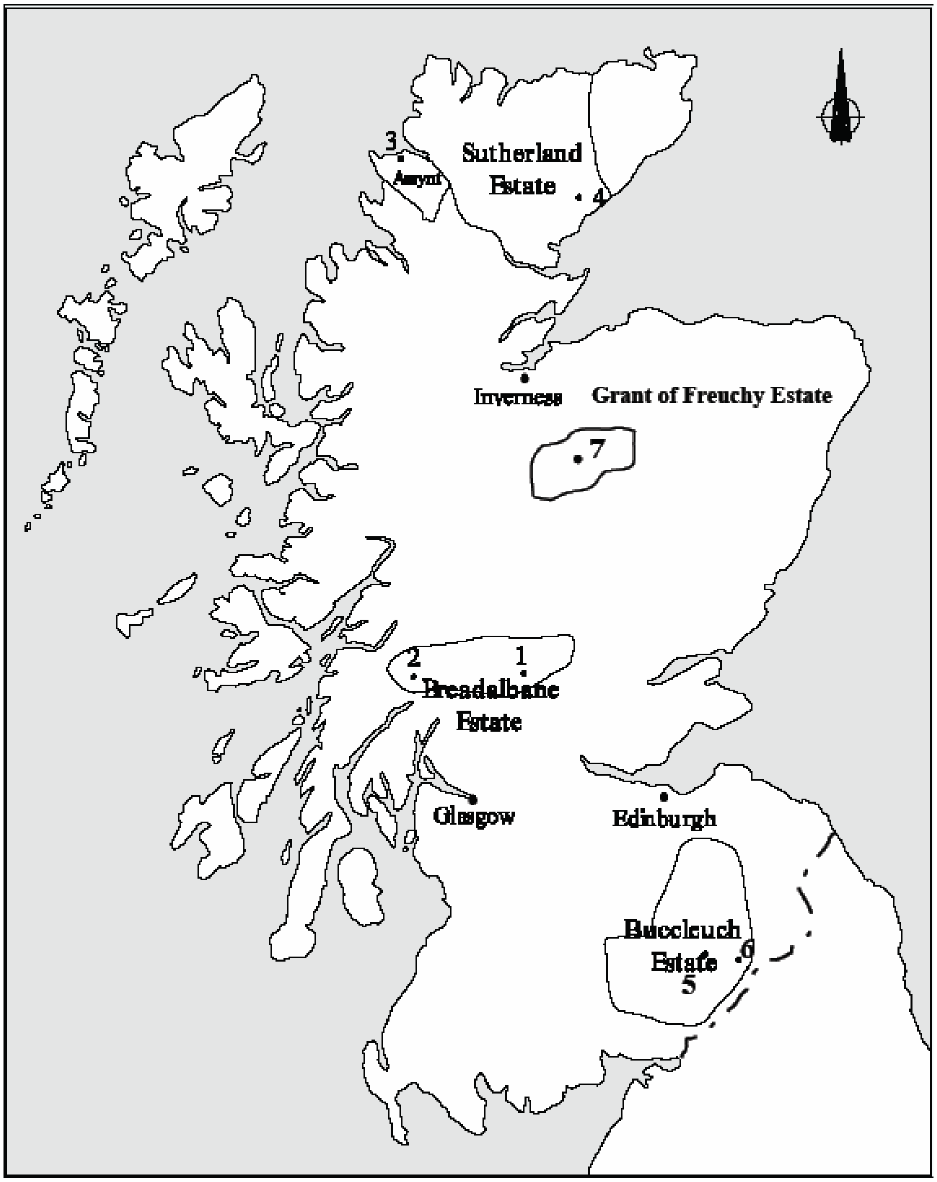

Figure 1. Locations of all sites investigated in project. *Breadalbane Estate*: (1) Leadour farm & shieling, Loch Tay, (2) Corries shieling, Glenorchy; *Sutherland Estate*: (3) Glenleraig farm & shieling, Assynt, (4) Rogart farm & shieling, Sutherland; *Buccleuch Estate*: (5) Bush of Ewes farm, Ewesdale, (6) Greenshiels shieling/farm, Liddesdale; *Grant of Freuchy Estate*: (7) Rynuie farm/shieling, Abernethy.

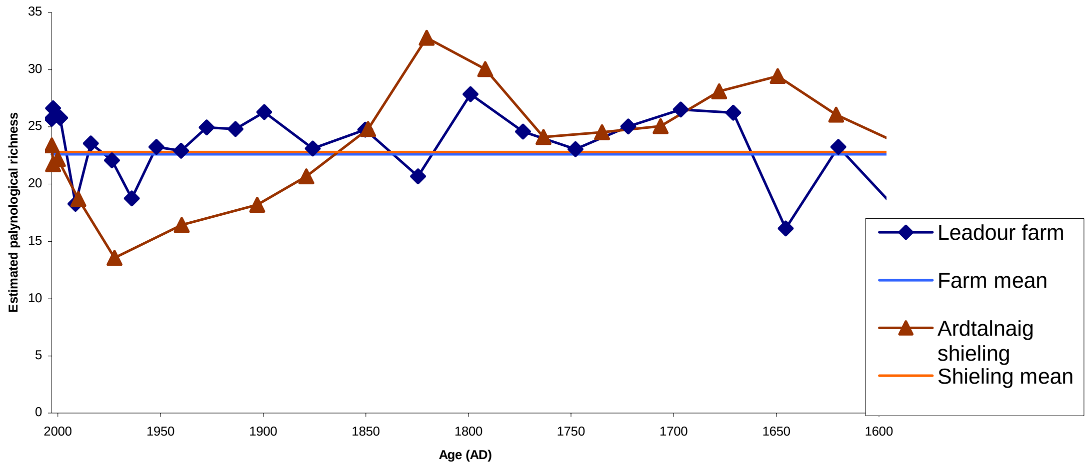

Figure 2a: Estimated pollen diversity score over time for two sites near Loch Tay, 1600 to present.

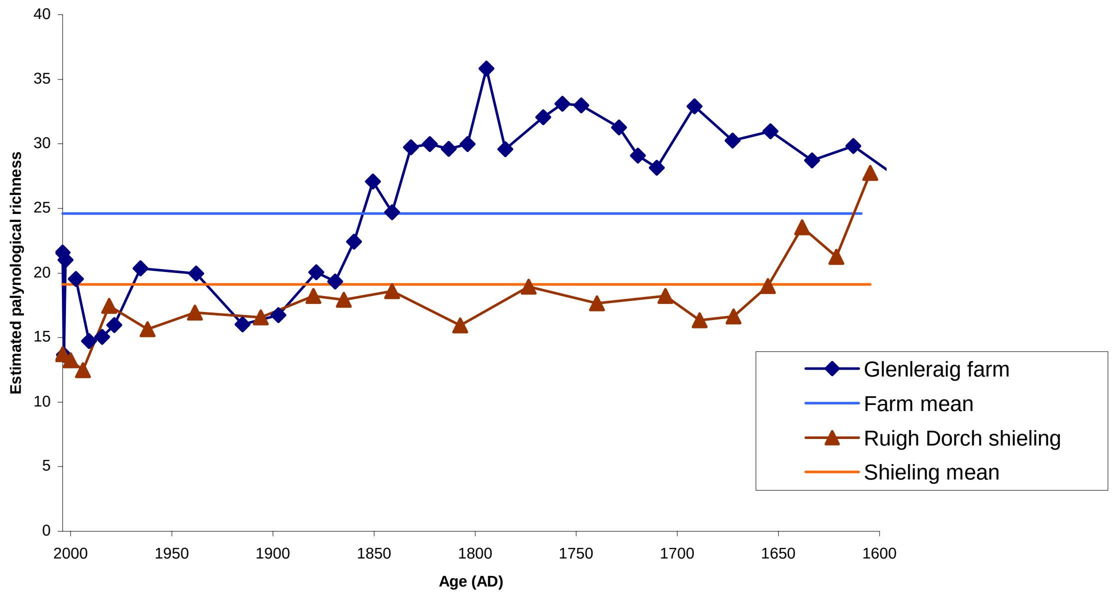

Figure 2b: Estimated pollen diversity score over time for two sites in Assynt, 1600 to present.

Table One – Variables used in model

Variable Name	Meaning	Main sources	**Type of data**
Dependent Variable:			
S_{it}	estimated species count at site *i* in year *t*	Pollen analysis	Continuous
Explanatory Variables:			
Lagged diversity	Species diversity estimate in previous 20 year period	Pollen analysis	Continuous
Management			
Site intensity	Intensity of use through year (5=year round; 1= abandoned)	Estate records	categorical
Size change	Property amalgamation or split	Estate records	Count per 20 yr period
Management change	Eg enclosures, draining	Estate records	Count
Animal issues 1	Disease	Estate records	Yes/no
Animal issues 2	New breeds introduced	Estate records	Yes/no
Prices			
Oats	Regional market price	Fiars data, estate records	In £/240
Bere (barley)	Regional market price	Fiars data, estate records	In £/240
Sheep	Regional market price	Estate records; Royal Highland Agricultural Society,	In £/240
Cattle	Regional market price	Estate records; Royal Highland Agricultural Society	In £/240
Wages	Labourer's wages	various	In £/240
Tenure			
Owner change	Change in ownership of site	Estate records	Yes/no
Occupier change	Change in tenant farmer	Estate records	Yes/no
Environmental			
Temperature	Mean monthly	English data	Degrees C
Rainfall	total annual	English data	mm
Extreme weather events	Storms, floods unusual enough to be recorded.	Estate records	Count
Other			
Extreme civil events	War, disease, famine etc	Estate records	Count
Harvest Failure	Failure of crops (due to disease or climatic conditions etc.)	Historical Records	Yes/No
Demand Drivers			
Act of Union	Union between Scotland and England,	Historical Fact	Yes/No
Refrigeration	Introduction of refrigerated transport.	Historical Fact	Yes/No
English Population	Population of England	Pre 1800: expert opinion Post 1800: Census	Count
Scottish Population	Domestic Population	Pre 1800: expert opinion Post 1800: Census	Count
War	Major conflict during period	Historical Fact	Yes/No

TABLE Two: The effect of economic activity on biodiversity

Dep. variable: ***Biodiversity index (B_t)***	***(1)*** **2SLS**	***(2)*** **Fuller-LIML**	***(3)*** **2SLS**	***(4)*** **Fuller-LIML**	***(5)*** **Fuller-LIML**	**(6) Fuller-LIML**
B_{t-1}	0.571**	0.571**	0.573**	0.574**	0.579**	0.583**
	(3.93)	(3.87)	(3.90)	(3.83)	(3.91)	(3.89)
pcattle	-0.006**	-0.006**	-	-	-0.006**	-
	(-2.16)	(-2.16)			(-2.06)	
psheep	-	-	-0.07**	-0.07**	-	-0.07*
			(-2.08)	(-2.07)		(-1.98)
sizechange	0.607	0.607	0.609	0.610	-	-
	(0.53)	(0.53)	(0.53)	(0.54)		
mgtchange	-0.092	-0.092	-0.079	-0.079	-	-
	(-0.41)	(-0.41)	(-0.35)	(-0.35)		
andisease	-0.605	-0.607	-0.583	-0.585	-	-
	(-0.55)	(-0.55)	(-0.53)	(-0.53)		
annewbread	1.583	1.582	1.662	1.661	1.296	1.388
	(1.12)	(1.12)	(1.17)	(1.17)	(1.05)	(1.12)
mgtinten	0.515**	0.514**	0.505**	0.504**	0.504**	0.493**
	(2.31)	(2.30)	(2.25)	(2.24)	(2.33)	(2.26)
extrweather	-0.229	-0.229	-0.222	-0.222	-	-
	(-0.63)	(-0.63)	(-0.60)	(-0.61)		
extrcivil	-0.095	-0.094	-0.118	-0.117	-	-
	(-0.13)	(-0.13)	(-0.17)	(-0.17)		
constant	8.563*	8.565*	8.546*	8.549*	8.293*	8.238*
	(1.89)	(1.87)	(1.85)	(1.82)	(1.80)	(1.75)
Sargan over-identification test	$\chi^2_{(4)} = 2.504$	$\chi^2_{(4)} = 2.504$	$\chi^2_{(4)} = 2.841$	$\chi^2_{(4)} = 2.840$	$\chi^2_{(4)} = 1.416$	$\chi^2_{(4)} = 1.732$
	(0.643)	(0.643)	(0.584)	(0.584)	(0.841)	(0.784)
Anderson canonical	$\chi^2_{(5)} = 38.44$		$\chi^2_{(5)} = 38.07$		$\chi^2_{(5)} = 37.10$	$\chi^2_{(5)} = 36.88$

correlations	(0.000)		(0.000)		(0.000)	(0.000)
Cragg-Donald under-identification	$\chi^2_{(5)} = 45.38$ (0.000)		$\chi^2_{(5)} = 44.86$ (0.000)		$\chi^2_{(5)} = 43.54$ (0.000)	$\chi^2_{(5)} = 43.23$ (0.000)
Stock-Yogo weak identification	6.04 (9.48)	6.04 (5.34)	5.97 (9.48)	5.97 (5.34)	6.10 (5.34)	6.05 (5.34)
First-stage F (B_{t-1})	$F(6,95) = 8.99$ (0.000)		$F(6,95) = 8.99$ (0.000)		$F(6,100) = 9.37$ (0.000)	$F(6,100) = 9.37$ (0.000)
First-stage F (*pcattle/psheep*)	$F(6,95) = 112.46$ (0.000)		$F(6,95) = 111.70$ (0.000)		$F(6,100) = 114.8$ (0.000)	$F(6,100) = 114.75$ (0.000)
Shea partial R^2 (B_{t-1})	0.276		0.274		0.268	0.267
Shea partial R^2 (*pcattle/psheep*)	0.669		0.663		0.651	0.647
SW normality test	0.992 (0.751)	0.992 (0.753)	0.992 (0.737)	0.992 (0.740)	0.993 (0.846)	0.993 (0.823)

Notes:1. There are 119 observations. All regressions include dummies for each site. 2. The instruments used are B_{t-2}, popenglish, garrison, union, refrigeration, pbere. 3. t-ratios are shown in parentheses below the estimated coefficients. An asterisk denotes significance at the 10% level and two asterisks at the 5% level. 2. LIML is Fuller's (1977) modified LIML with a=1. 5. The Sargan test is a test of overidentifying restrictions. Under the null, the test statistic is distributed as chi-squared in the number of overidentifying restrictions (the p-value is reported in parenthesis). 6. The Anderson (1984) canonical correlations is a likelihood-ratio test of whether the equation is identified. The Cragg and Donald (1993) test statistic is also a chi-squared test of whether the equation is identified. Under the null of underidentification, the statistics are distributed as chi-squared with degrees of freedom=(L-K+1) where L=number of instruments (included + excluded) and K is the number of regressors (the p-values are reported in parentheses). 7. The Stock and Yogo statistic is used to test for the presence of weak instruments (i.e., that the equation is only weakly identified). The critical value for a 10% bias in 2SLS is reported in parentheses (see Stock and Yogo (2005) for a tabulation of critical values). 8. The 1st stage F-statistic tests the hypothesis that the coefficients on all the excluded instruments are zero in the 1st stage regression of the endogenous regressor on all instruments (the p-value is reported in parenthesis). 9. Shea's (1997) "partial R-squared" is a measure of instrument relevance that takes into account intercorrelations among instruments. 10. The SW is the Shapiro and Wilk (1965) test for normality, for the residuals of the structural equation. The p-value of the test is reported in parentheses.