The Price of Land in Eastern Prussia: Data from Capital Gains Tax Records, 1891-1907

Scott M. Eddie*

Abstract: Using data on 438 individual sales of farm properties of 5 hectares or more in size, this article finds that simple area in hectares of the individual land types and variables for trend, sales to family members, and overall size of the property, explain nearly all the variation in the prices (both actual and »net«, i.e., adjusted for certain qualitative variables, deflated by an overall price index to 1913 marks). Buyers received a discount of about 21 per cent if the seller was a relative, but paid a premium of about 19 per cent to acquire a large property (over 100 hectares), contrary to the conventional wisdom of the time (1891 - 1907). The extra premium for a large property is consistent with a number of hypotheses, including that of owners associating in a »club« to exclude lesser beings on the basis of price and/or there being a consumption value for large properties but not for small ones.

In an agrarian society, land is the most important component of the national wealth, and the source of greatest status and security. Economic historians are often faced with an insuperable problem when working on such societies, however, because land is also the most difficult major component of national wealth for which to estimate a value. We know the landlords were rich, but *how rich* were they? The difficulty of estimation arises partly because (a) land has many uses for which both the quantity of output and the market price are difficult to estimate,[1] and (b) often the non-land inputs into the processes which

*** Address all communications to Scott M. Eddie, Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada M5S 3G7.** **The author would like to thank the University of Toronto for financial support and Professor Heinz Reif (Technical University, Berlin) and John Komlos (Ludwig-Maximilians University, Munich) for providing facilities which furthered the work on this project, and Dr. Christa Kouschil, Berlin, for helping me find and interpret the data on which this study is based.**

[1] Fox hunting springs immediately to mind, but there are many others.

produce these outputs are difficult to identify or to value, making a capitalization of the stream of net income from land an exercise in guesswork at best.[2]

Under such conditions, it would therefore be preferable to work directly from historical data on market prices of land. But land prices are among the most difficult price data to obtain, often because of legal restrictions on their publication.[3] Some have indeed been published - scattered data can be found for some areas of Prussia in local newspapers, for example - but finding enough of these data to allow statistical processing with any degree of confidence founders on either paucity of data or too high cost of collection.[4]

But where governments tax, they have an incentive to gather accurate data. In 1893, Prussia introduced a capital gains tax *(Wertzuwachssteuer,* also called the *Ergänzungssteuer).* Since it was to apply to land as well as other assets, the land registry offices *(Katasterämter),* which also had responsibility for the levying of the land tax, were required to keep records of the sales of landed properties in their districts. One incomplete set of these records, from the riding *(Kreis)* of Angerburg in East Prussia, has survived in the Prussian State Archive (GStPK, Rep. I 151II) Angerburg (today Wêgorzewo in Poland) was a landlocked riding southeast of Königsberg (today Kaliningrad); the *Kreisstadt*[5] Angerburg is located at about 21°45' East longitude and about 54° 12' North latitude, almost exactly 100 kilometres southeast of Königsberg.

These records from Angerburg are the so-called *Urlisten* (fundamental lists), which list the details of individual sales of so-called »Group B« properties, farmsteads with buildings *(Hofbesitz);* from them I have taken the data on all sales of properties of 5 hectares or more,[6] a total of 438 usable observations. The time distribution of these observations is as follows:

[2] In a country such as England, where hunting rights can be sold separately from the land, or in Prussia, where mineral rights and land ownership did not necessarily go together, there sometimes exist data which one could use to estimate non-agricultural, non-forestry values of land. These are, I think, rather the exception than the rule.

[3] The Prussian Statistical Office published extensive summary data on the price of land, based on the price data collected for the capital gains tax (PSL 1917). Since the Prussian statisticians were very careful to make it impossible to connect a price to any given property or even any given locality, these interesting data are useless for estimating the value of land anywhere in Prussia.

[4] Data on land purchases by a government agency are sometimes available, for example in reports to a parliament, and in an earlier paper (Eddie 1993) I used such data from a Prussian government agency, the Settlement Commission for Posen and West Prussia *(Ansiedlungskommission für Posen und West Preußen).* The data used in the present paper have the great advantage of being data on private transactions.

[5] Seat of the government of the riding.

[6] The lower limit was chosen to limit the data set to »genuinely« agricultural properties, based on a notion common in the German literature that units under 5 hectares were not big enough to support a family. See for example Sering 1910, esp. pp. 30-33. Cost of data collection also contributed to limiting the data set in this way.

1891	1892 1	1893 5	1894	1895	1896	1897	1898	1899
1			66	64	49	43	55	28
1900	1901 0	1902 0	1903 0	1904	1905	1906	1907 2	
27				22	37	38		

The Prussian Finance Ministry printed forms for the *Urlisten,* on which the *Katasterämter* were to record the relevant data: the location and designation of the property, the name of the purchaser, the price, whether the purchaser and seller were related to each other, if it was a forced sale, the data on types of land and taxable value of the property and its buildings, and the condition of cultivation, of any standing timber, of the buildings, and of the movable inventory (stock, seed, machinery, etc.). All the important variables which the Finance Ministry thought could influence the selling price were therefore to be set down on these forms.[7] The price data were reported to the Katasteramt from the local court office *(Amtsgericht)* (PSL 1917, p. i), and of course the data for amounts of different types of land and their respective GSREs, along with the taxable value of buildings, were on file at the land registry office. Where the information on condition of cultivation, etc. came from is less clear: From occasional written comments on the forms, it appears that registry officials may in some cases have actually looked at the property. In any case, Angerburg riding was a small place, and the land registry officials could be presumed to have had an intimate knowledge of the properties in their bailiwick.

The net yield of the land for the land tax *(Grundsteuerreinertrag,* hereafter GSRE) was established in the early 1860s,[8] and measured not what the land actually produced, but what it could produce under normal management and local market conditions. The GSRE was thus a measure of the productivity of land; carried out fairly and accurately, its assessment should have produced an extremely good proxy for market value of land. Each plot of land on a property was estimated separately according to its size, land type, and quality, and the sum of these taxable net yields was then the taxable value (GSRE) of the property as a whole. Each year the Prussian government set the percentage of this value which was to be the land tax payable; that percentage was the same throughout the country. The forms included a calculation not only of price per

[7] There were also sections on the form to make adjustments to the actual purchase price - additions or subtractions - to arrive at the »net price« *(Reinpreis)* of the property, and room to write in the reasons for these adjustments. The reasons most often had to do with the price being too low (in one case too high) because the buyer and seller were related, or because of the state of cultivation, wood, buildings, or inventory. More about this later.

[8] Completed 1864.

hectare, but also of price per Thaler[9] of GSRE - both ratios were common measures of the unit value of land in Prussia at the time.

The tax reform in 1893 also fixed the percentage of the land tax permanently, and turned its revenue over to the communities. Surcharges could be added on top of this tax, a potentially complicating factor in the analysis of price formation. For the single riding, and over the relatively short time span, for which this paper has data, this was not a serious problem.

I

Since the Ministry of Finance form required that both the area and tax assessment of each type of land (plough, garden, meadow, pasture, woodland, moor, bodies of water, waste, and farmyard) be recorded, data on tax assessments are at our disposal. Were they current, it would of course have been theoretically preferable to use the tax assessment data to estimate the determinants of price, since it is the income potential of any asset that should be the prime determinant of its price. I have nevertheless used the *areas* of each type of land, rather than their tax assessments, as independent variables in the regressions which follow, both because the tax assessment data were not current and because I have an interest in using the analysis and its results with other available data for Prussia.[10] These other data often have the areas of each type of land, but never the tax assessments by land type.[11]

The data on land area might not have been current either, because the Prussian land tax did not change its assessments of land, precisely in order to give property owners an incentive to convert from low-value to high-value uses (Lesigang 1900, p. 898). But these conversions occurred slowly, because costly, and required a confidence in the future course of prices which often did not exist because of frequent policy changes and foreign competition, not only for grains but also especially for wool, in the nineteenth century.

[9] 1 Thaler - 3 marks.

[10] In this I am also following the lead of the Prussian statistical office, which considered the price per hectare a much more important indicator than the price per unit of GSRE: »Daß [d]er [Grundsteuerreinertrag] seit 1861 im großen ganzen in starrer Unveränderlichkeit verblieben, daher gewissermaßen überlebt ist und nicht mehr durchweg den tatsächlichen Verhältnissen entspricht, beeinträchtigt natürlich seinen Wert als volkswirtschaftlichen Gradmesser. Demnach ist insbesondere der Kaufpreis pro *ha* viel wichtiger zu nehmen als der für 1 *M* Grundsteuerreinertrag.« (PSL 1917, p. ii)

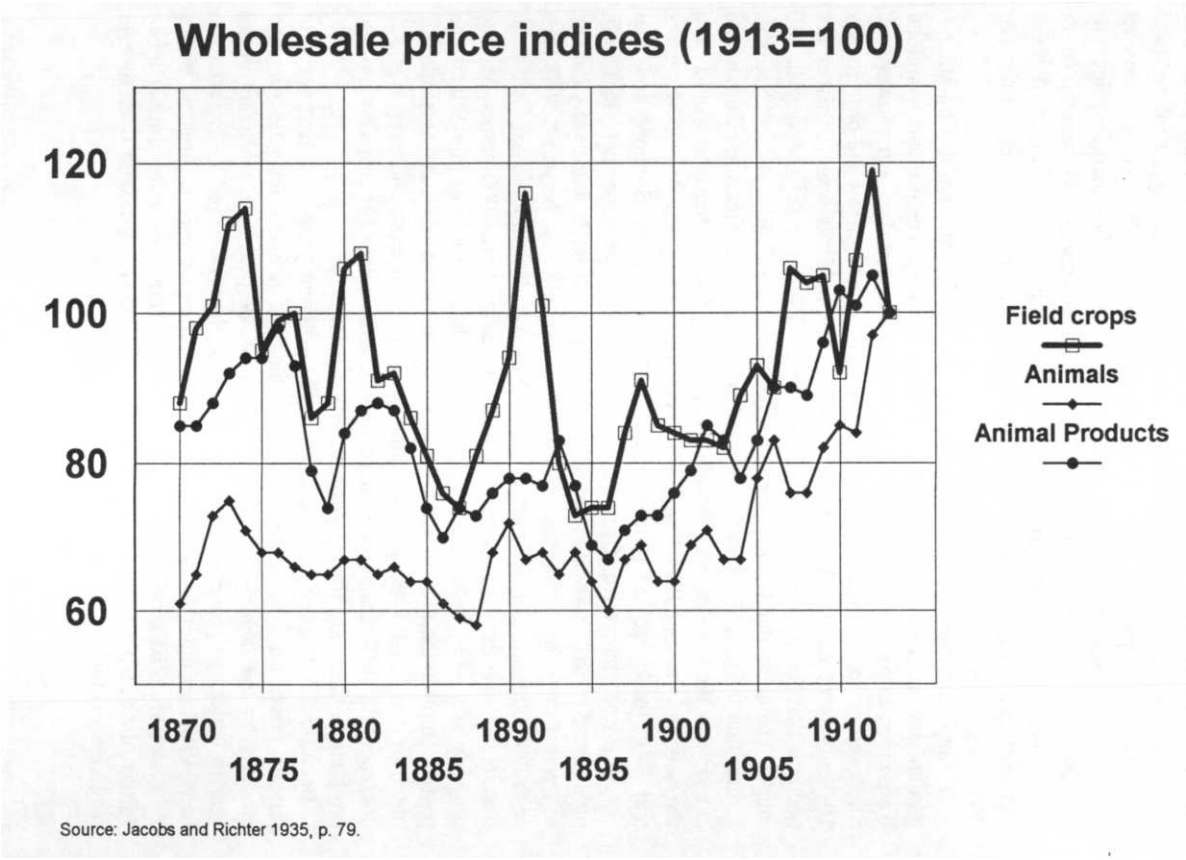
[11] Because the value assessment for the land tax was carried out in the early 1860s and completed in 1864, the relative tax assessments for the various types of land might no longer have represented their relative income potentials at the time these data were gathered. To check this assumption, I also ran the regressions using the tax assessments; the results were marginally less good than those using simple area, an additional justification for the use of an areas in the regression equation.

The expected income from land was much more affected by volatility and trends in output prices than by conversions of land use from one type to another. Because of differential price changes as between field crops and animal products, the relative market values of different land types (ploughland vs. meadow and pasture) could be expected to change. Their tax assessments did not, however, so that divergences between relative market prices of different land types and their relative tax assessments could have become larger than divergences between relative market prices and relative areas of these different land types, since different land types had different tax values per hectare.

A simple hypothetical example should help clarify this point Ploughland, garden, and meadow had the highest GSREs, those for pasture and woodland were much lower, and those for bodies of water and moorland even lower still. Suppose a property were made up of equal parts of ploughland and pasture, say 10 hectares of each, and that the GSRE for ploughland was 10 marks per hectare, that for pasture 4 marks, so the total GSRE for the property is 140 marks. The areas stand in a 1:1 ratio to each other, the GSREs at 10:4. Suppose further that the GSRE estimation was a fair and accurate measure of the actual net yield at the time, so that initially the market prices of the land would have the same ratio as the GSRE, 10:4. Let us take these two prices to be 5000 marks and 2000 marks, so the property was worth 7000 marks. Suppose now that output prices change: (a) Crop prices fall, reducing the actual net yield from ploughland by 20%. Abstracting from any inflation or deflation, or any change in interest rates, the 10 hectares of ploughland should now be worth 20% less, or only 4000 marks, (b) Animal product prices rise, doubling the net yield from pasture, so now the 10 hectares of pasture are also worth 4000 marks, and the property as a whole 8000 marks. At first the relative prices of ploughland and pasture stood in the ratio of 10:4, the same as the ratio of their tax assessments, but now they stand at 1:1, the same ratio as their areas. Charts 1 and 2 document the differential paths taken by wholesale prices for crops and animal products in Germany for the period 1870 to 1913.

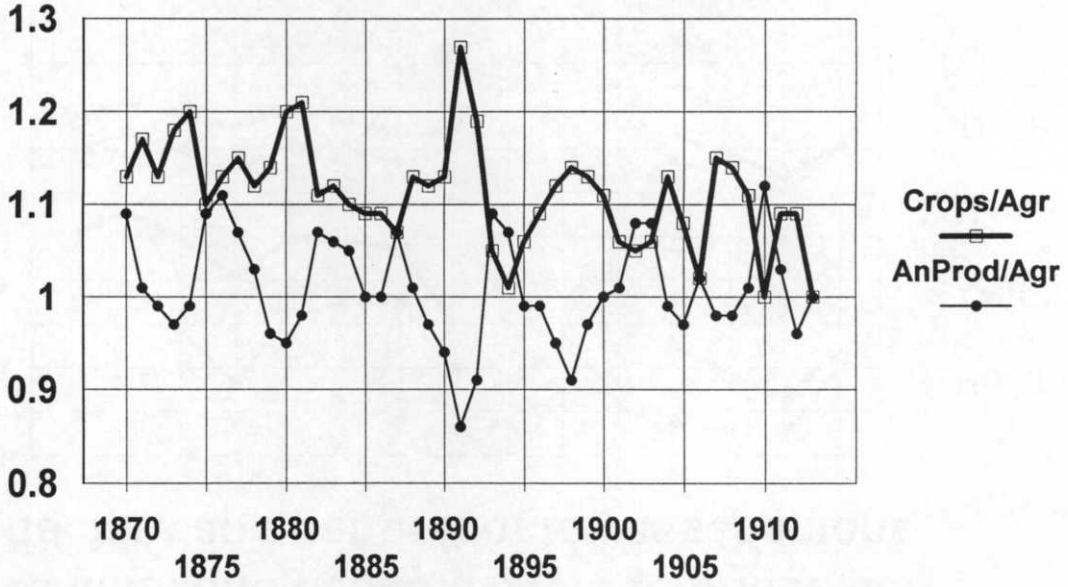
Because the data are for property sales in different years, one first needs to put the prices on a common basis: I deflated the actual price paid for each property by the Jacobs-Richter index of wholesale prices for Germany in the year in which the property was sold (Jacobs-Richter 1935, p. 79). Because the Jacobs-Richter index is based on 1913=100, deflation converts purchase prices to marks of 1913 purchasing power.[12] Chart 3 shows how the annual mean values of price per hectare and price per unit of GSRE changed over time in Angerburg riding.

[12] **Note that this is not the same thing as what the property would have sold for in** 1913. **That latter figure would have to be calculated from the coefficients in the regression equation. For a discussion of this point, see Eddie** 1993.

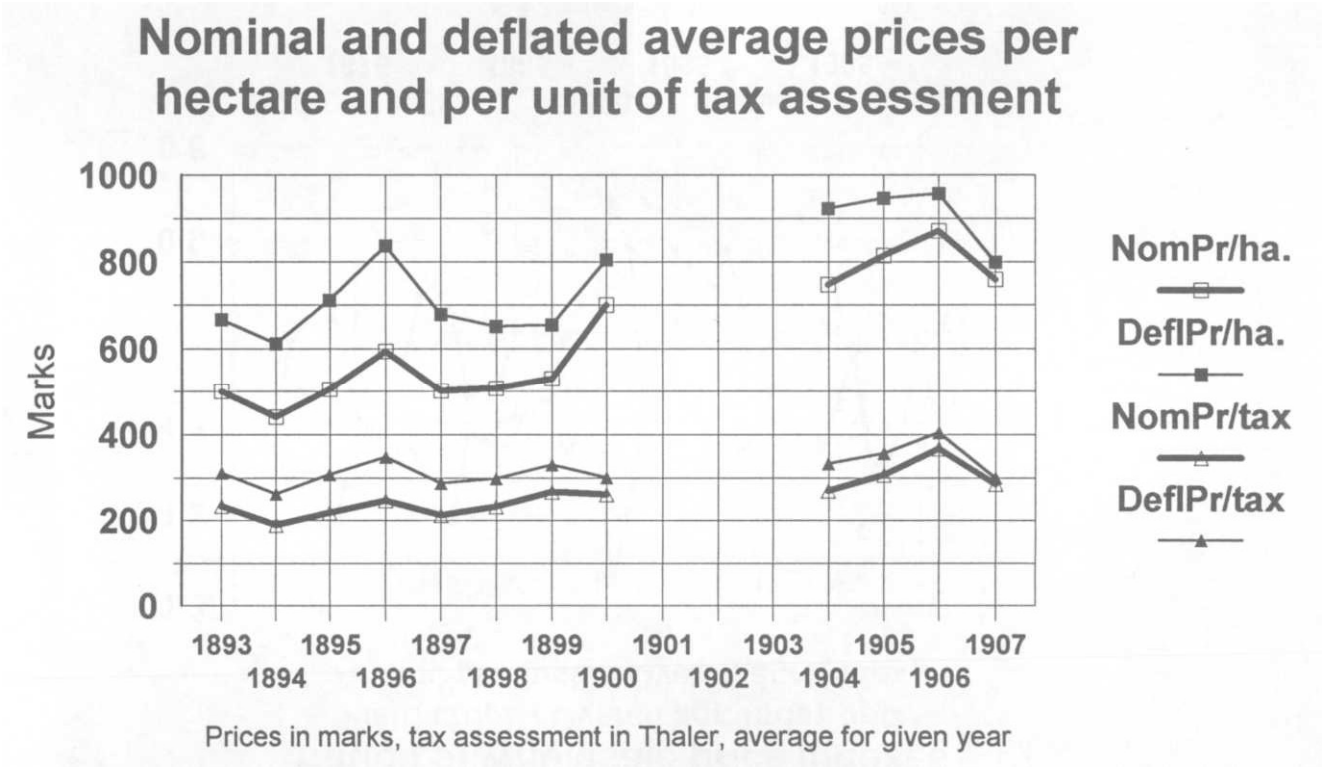

Wholesale price indices (1913=100)
120
100
80
60
1870
1875
1880
1885
1890
1895
1900
1905
1910
Field crops
Animals
Animal Products
Source: Jacobs and Richter 1935, p. 79.

Ratios of wholesale price indexes

Field crops / Overall agr. index and Animal products / Overall agr. index



Source: Jacobs and Richter 1935, p. 79.
Note: Overall agr. price index includes live animals (not included in animal products index)


Nominal and deflated average prices per hectare and per unit of tax assessment
Marks
1000
800
600
400
200
0
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
NomPr/ha.
DeflPr/ha.
NomPr/tax
DeflPr/tax
Prices in marks, tax assessment in Thaler, average for given year

II

As a first estimation, I regressed the deflated purchase price against the number of hectares of each type of land, a trend factor, and variables to test (a) the proposition that properties larger than 100 hectares (the universally accepted dividing line between large and smaller properties in Prussia at the time) sold for either more or less per hectare than smaller properties, *ceteris paribus,* and (b) that transactions among relatives did not take place at the same prices as arm's-length transactions. The trend factor is simply the quantity (year minus 1891) times the total number of hectares in the property. It thus begins with a value of 0, and its coefficient represents the annual increase in the price of landed property because of influences not captured in the regression equation, in marks of 1913 purchasing power per hectare per year. The »big property« variable is the result of assigning a dummy variable a value of 1 if the property was larger than 100 hectares, 0 otherwise, and multiplying the value of this dummy variable by the total number of hectares in the property. Its coefficient represents the premium paid to acquire a larger property, in 1913 marks per hectare. If a transaction occurred between relatives, a dummy variable received a value of 1, if not, of 0. This dummy was then multiplied by the total number of hectares in the property. The coefficient would therefore be the discount (or premium) given to a relative in the sale of a property, in 1913 marks per hectare. The results of this regression are listed in Table 1.

The estimation produces positive and highly significant coefficients for all the directly productive land types: ploughland, garden, meadow, pasture, and woodland. Bodies of water or the size of the farmyard seem not to have affected the price, for their coefficients are insignificant, as are those for moorland and waste.[13] A small but significant time trend of some four 1913 marks per hectare per year can be noticed in the price, and buyers appear to have received a discount of 163 marks per hectare from a seller who was a relative. This was about 21 per cent of the average price per hectare. It further appears that larger properties may have commanded premium prices, since the coefficient for the »big property« variable is positive and statistically significant. I will return to this important point below.

Were properties sold within the family different from the others? Not really, except that they were neither so heavily concentrated among the smallest properties, nor among the largest, as were the rest of the properties:

[13] t-test; minimum 90% level of confidence.

Number of sales to:	Overall size of property in hectares 5-20	20-50	50-100	100-200	200-500	500-1000	Over 1000	∑
Family	46	45	22	3	4	0	1	121
Non-family	166	63	38	30	18	2	0	317
Total	212	108	60	33	22	2	1	438

Table 1: Initial regression

Regression with deflated purchase price as the dependent variable
Mean = 40401 Marks per property, Standard deviation = 70272
Marks Weighted mean price per hectare = 762 Marks

Independent variable (unit of measurement is hectares unless otherwise noted)	*Coefficient*	*Std. Error*
Ploughland	333	48.9
Garden	1813	627
Meadow	808	101
Pasture	578	80.9
Woodland	317	123
Water	-71.9	74.9
Moor	-1251	1046
Waste	-2771	1903
Farmyard	3232	3221
Building value (marks)	4.2	5.30
Trend*	4.13	0.490
Property over 100 ha.*	150	27.8
Relative*	-163	24.4
Constant term / standard error of estimate:	-1168	12678
R^2	0.968	
Number of observations / degrees of freedom	438	424

* See text for explanation of these variables.

There seems no reason to suspect that there is anything unusual about the value of the coefficient for the »family discount«. That the constant term is insignificant is in fact comforting: Since we have included all types of land in the estimating equation, a significant constant term would represent a payment (or a deduction) for some sort of intangible that we have been unable to take into account.[14] The estimated value of the coefficients for ploughland, garden, meadow, pasture, and woodland are very different and - except for ploughland - they also stand in rank order of average assessment,[15] as we should expect. The relative sizes of the coefficients and the relative sizes of the average assessments are very different, however, as the comparison below shows:

	PLOU	Gar-den	Mead-ow	Pas-ture	Wood-land
a. Coefficient	333	1813	808	578	317
b. Average GSRE/hectare	2.70	4.27	3.18	0.83	0.62
Ratio a / b	123	425	254	697	512

The relatively small coefficient for ploughland, compared to meadow and pasture, is likely the result of the differential price trends in the chief outputs of each (charts 1 and 2), which favoured the production of animal products compared to field crops. It may also be directly related to the relatively high coefficients for pasture and woodland in another way: As previously mentioned, the Prussian land tax system deliberately did not change assessments when conversions from one type of land use to another were made. The pattern of coefficients observed is consistent with such an unrecorded conversion of pasture and woodland to ploughland. From the point of view of accuracy of our estimates, the results of such a policy are unfortunate, but there is nothing we can do about it. The amount of land thus converted to ploughland is likely to have been small, given that the prices of field crops were so volatile and trended down to at least the mid-1890s (chart 1).

One might expect that larger properties would sell at a lower price per hectare, *ceteris paribus,* because of transactions cost in dividing them into smaller units for sale to several buyers, or because of imperfections in the market for complementary inputs, labour and credit likely being the most important.[16] The results here seem to indicate exactly the opposite, that a

[14] This reinforces the same result from my earlier study of land prices (Eddie 1993).
[15] The rank order is garden, then ploughland and meadow with about the same assessment, pasture, woodland.
[16] If there were economies of scale in agriculture, that might be another reason for a

relatively substantial premium was paid for larger properties. This would be consistent with large properties having a consumption value in addition to their value as a productive asset,[17] or with a »clubbing together« of their owners to exclude lesser mortals on the basis of price.[18]

III

The unusually good fit of the estimated regression equation suggests that the additional information about quality of cultivation and the like will not likely add much to the explanatory power of the equation. One should nevertheless use this information, if only to see if it makes any appreciable difference in the previous estimates. Accordingly, several new variables were added:

1. Variables describing the state or condition of the property. On the forms the qualitative judgments were supposed to be »good«, »medium«, »bad«, or »missing«, but these notations were not always adhered to. **I** have taken »mediocre« *(mäßig)* and »very mediocre« *(sehr mäßig)* to mean »bad« in assigning values to the dummy variables described below. Under »condition of standing timber«, a blank meant there was no woodland on the property; in other cases where the notation was left blank, **I** examined each observation individually to interpret what the blank could mean (usually »medium«). Other unusual notations are explained below under the relevant variable.

   a. Good or bad cultivation: two dummies, set =1 if the cultivation condition was listed as »good« or »bad«, respectively, = 0 otherwise. Each was then multiplied by the sum of cultivated land on the property (ploughland plus garden plus meadow). *The coefficients would then be the premium resp. discount on the price of the property as a result of good or bad state of cultivation, in 1913 marks per hectare of the cultivated land.*

   b. Good or bad buildings: two dummies, set = 1 if the condition of the buildings was listed as »good« or »bad«, respectively, - 0 otherwise. Each was then multiplied by the tax assessment on the buildings. *The*

premium price paid for larger properties. The introduction to the Prussian land prices volume asserts that prices per hectare were higher for small properties because there was a much stronger demand for them, without going into the questions of why this demand should be stronger, or why larger properties were not divided into smaller parcels to take advantage of this stronger demand (PSL 1917, p. ii).

[17] Such was the case in Prussia: Many large properties were *Rittergüter* (knightly estates) which entitled their owner to a seat in the local assembly *(Kreistag).* Unfortunately the data used here do not permit me to identify precisely which of the properties were *Rittergüter.*

[18] In an as-yet unpublished paper, Peter B. Morgan presents an elegant exposition of the theory behind such »clubs« (Morgan 1997).

coefficients would then be the premium resp. discount on the price of the property as a result of good or bad state of the buildings, in 1913 marks per mark of assessed value of buildings.

c. Good or bad stand of wood: The »good« dummy was set = 1 if the wood stand was listed as good, - 0 otherwise. The »bad« dummy was set = 1 if the wood stand was listed as »bad«, »burned«, »cut down«, or »missing«, = 0 otherwise. These dummies were then multiplied by the number of hectares of woodland on the property. *The coefficients would then be the premium resp. discount on the price of the property as a result of good or bad state of the standing timber, in 1913 marks per hectare of woodland.*

d. Good or bad movable inventory: The »good« dummy was set = 1 if the condition of the movable inventory was listed as »good« or »very good«, = 0 otherwise. The »poor« dummy was set - 1 if the condition was listed as »poor«, »very poor«, »quite bad«, or »missing«, - 0 otherwise. These dummies were then multiplied by the size of the property in hectares, on the ground that the inventory should be roughly proportional to the size of the property. *The coefficients would then be the premium resp. discount on the price of the property as a result of good or bad state of the movable inventory, in 1913 marks per hectare.*

2. If there was a business on the property, such as an inn (most common), brickworks, or distillery, a dummy variable was set = 1; if not, = 0. *The coefficient of this variable would then show the mean value of these businesses, in marks.*

The results of the second regression are shown in Table 2.

The results from this new regression equation are mixed at best: First, the fit is only marginally improved; the R^2 has improved by only one percentage point. Given the loss of degrees of freedom, that is no net gain. Moreover, the additional variables make no fundamental difference in the estimates shown in Table 1: While the coefficients for plough, garden, meadow, pasture and woodland may at first glance appear to stand in a more »reasonable« relation to each other, in fact the spread between the ratios of the coefficients to the average GSRE per hectare has increased:[19]

[19] **Since I took the different quality assessments to apply to different types of land (cultivation to plough, garden, and meadow; wood to woodland, inventory to all land) maybe this was to be expected. My thanks to John Komlos for pointing this out to me.**

Table 2: Regression including qualitative variables

Regression with deflated purchase price as the dependent variable
Mean = 40401 Marks per property, Standard deviation 70272 Marks
Weighted mean price per hectare = 762 Marks

Independent variable (unit of measurement is hectares unless otherwise noted)	Coefficient	Std. Error
Plough land	279	43.3
Garden	676	538
Meadow	672	90.6
Pasture	468	72.0
Woodland	485	130
Water	24.4	65.0
Moor	-609	933
Waste	769	1933
Farmyard	1693	3210
Building value (marks)	45.2	6.08
Trend*	6.55	0.513
Property over 100 ha.*	118	23.9
Relative*	-103	21.8
Good cultivation*	-104	20.6
Bad cultivation*	-126	65.4
Good building (marks)*	21.7	4.58
Bad building (marks)*	20.3	9.19
Good wood*	66.4	201
Bad wood*	886	217
Good inventory*	131	17.0
Bad inventory*	-174	43.9
Dummy for existence of business (1 or 0)*	3701	2366
Constant term / standard error of estimate:	-350	10727
R^2	0.978	
Number of observations / degrees of freedom:	438	415

* See text for explanation of these variables.

From Table 2:	Plough	Gar-den	Mead-ow	Pas-ture	Wood-land
a. Coefficient	279	676	672	468	465
b. Average GSRE/hectare	2.70	4.27	3.18	0.83	0.62
Ratio a / b	103	158	211	564	751

From Table 1 :	Plough	Gar-den	Mead-ow	Pas-ture	Wood-land
a. Coefficient	333	1813	808	578	317
b. Average GSRE/hectare	2.70	4.27	3.18	0.83	0.62
Ratio a / b	123	425	254	696	511

The only pair of »condition« variables that works as expected is that for inventory. Buyers appear to have paid more if the inventory were good, less if it were bad. Perhaps this was so because inventory was more of a quantitative than a qualitative judgment: often the relevant comment was either »without« or »missing« in the »condition of inventory« blank on the form, rather than a qualitative statement. The others confined themselves almost exclusively to qualitative statements (although »missing« often appeared under »condition of wood« for properties which had no woodland; for those observations the variable would have received a value of 0 irrespective of the comment, since the dummy for condition was multiplied by the number of hectares of woodland).

The other pairs of quality variables - for cultivation, standing timber, and buildings - produce either insignificant or anomalous results. Since the officials who filled in the forms always estimated that good cultivation either added to the price or did not affect it, it makes no economic sense for it to have detracted from the price, as implied by the negative coefficient. The negative coefficient for bad cultivation was to be expected, but the positive one for bad buildings was not. Both of them should have detracted from the price, and so they were treated on the forms by those who recorded the data. Similarly with the condition of wood - maybe a good condition would not add much (and often it did not on the forms), but bad condition was recorded as having either no effect or a negative effect, so a positive coefficient for this variable is also an anomaly.

These results suggest there is either missing information, or some interaction among the variables which the specification of the regression equation does not capture. This latter problem, in the form of multicollinearity, became especially acute when I tried to introduce variables to take account of major policy changes that should have affected the price of land (the Caprivi treaties, the tariff of 1902) and the effects of changing prices and yields on expected revenue from land.[20] The regressions run using these variables are not reported here because they added nothing to the »fit« of the regression equation, produced no significant coefficients for the variables themselves, and sometimes led to increased standard errors of other coefficients with which they were intercorrelated.

In sum, the additional variables of a qualitative nature seem to add little if anything to the explanatory power of the basic regression; indeed, except perhaps only for the case of the existence of a business on the property, they seem to introduce further problems rather than to add information. That should not be particularly disturbing, however, since the original, simple formulation produced such excellent results.

IV

There is also another, indeed more direct, approach available for taking the qualitative variables into account. Since the government officials calculated for most properties a »net price« *(Reinpreis)* after adjustments for these quality variables and for such things as transactions among relatives, we can test our original results by using this net price directly. Since the officials made no explicit corrections for the passage of time or the size of the property, these variables must still be included in any regression using »net price« as the dependent variable. But since their consideration of the qualitative variables and whether or not there was a business on the property led to explicit adjustments in the price, these variables must be omitted from the estimating equation. We thus come to an even simpler formulation than that of the first regression (Table 4).

It seems that the officials in the *Katasteramt* in Angerburg had a very good knowledge of the land market in their area. The results are very close to those of the original regression analysis, and - in contrast to the first regression - the coefficient of every variable, with the exception only of size of the farmyard, is highly significant (95% level for water, 99% level or better for the others). They also stand quite close to the values estimated in the original equation.

These variables did have significant effects on the prices paid by the Settlement Commission *(Ansiedlungskommission)* for the provinces of Posen and West Prussia (Eddie 1993).

Table 3: Regression using adjusted prices

Regression with deflated "net price" as the dependent variable Mean = 40378 Marks per property, Standard deviation = 69134 Marks Weighted mean price per hectare = 788 Marks

Independent variable (unit of measurement is hectares unless otherwise noted)	Coefficient	Std. Error
Ploughland	433	49.1
Garden	1731	576
Meadow	806	96.5
Pasture	655	76.1
Woodland	307	121
Water	471	227
Moor	-5770	1283
Waste	-5884	1786
Farmyard	5147	3053
Building value (marks)	23.6	7.47
Trend*	3.74	0.446
Property over 100 ha.*	153	26.8
Constant term / standard error of estimate:	230	11613
R^2	0.973	
Number of observations / degrees of freedom:	401	388

* See text for explanation of these variables.

That the coefficients of the main land types are close to each other is very reassuring for the quality of the original estimating equation. Of particular note is the essential identity of the value for a large property in the two equations. The claim that large properties would sell for less per hectare was not true, neither on average nor at the margin, in the riding of Angerburg. Among the 401 properties for which a »net price« was calculated, the weighted average actual deflated selling price[21] per hectare for properties over 100 hectares was

[21] **Sum of all selling prices in 1913 marks divided by the sum of all areas of the properties in question.**

845 marks vs. 698 marks for properties under 100 hectares. The »net prices« were 841 and 749 marks, respectively. For the entire data set, the differences in the averages are much smaller, 787 vs. 729 marks, suggesting that the sources of difference in the two equations lie primarily with the 37 properties for which the officials at the land registry office did not see fit to calculate a »net price«.

This latter consideration suggests that the original equation, using deflated actual selling price as the dependent variable, should be run on the reduced data set, to see if the properties which the officials deemed somehow unsuitable for a calculation of »net price« were a group of outliers that would substantially affect the results of the estimation. The results of the original regression equation, but using this reduced set, are compared to the other results in Table 4.

That the results on the reduced set should be so essentially identical suggests either (a) that the adjustments made by the land registry officials were, on the whole, marginal, or that they followed an essentially rule-driven procedure[22] which is well simulated in a simple regression; or (b) that the regression used has basically captured the essence of their adjustments, however complex they may have been. Either way, this gives very strong support to the formulation here used to estimate the price of landed properties in the riding of Angerburg.

V

The foregoing considerations bring us to the final form of the estimating equation to be used here on the full data set: it is simply the equation of Table 1 with the addition of a dummy variable for the existence of a business on the property. The results of this estimation are presented in Table 5.

Including the existence of a business into the final form of the equation is, of course, logical, although in fact it had no appreciable effect on the overall statistical outcome, perhaps because a business was found on only 34 of the 438 properties. A more likely reason, however, is that there was considerable variety in the value of the businesses, but the dummy variable measured only presence or absence. This crude formulation produced a coefficient significant at the 90% level: if one is willing to accept that level of confidence, the average business appeared to have been worth about 5000 marks of 1913 purchasing power.[23] Otherwise the signs and values of the rest of the coefficients did not change much.

In this final form we find the discount for a family member of 161 marks of 1913 purchasing power per hectare, still 21% of the average price per hectare of the properties in this data set. The premium for a large property was 145

[22] They were, after all, bureaucrats.

[23] Breaking the businesses down by type and assigning a dummy for each did not offer any improvement, hence no further elaboration of the business dummy variable.

Table 4: Comparison of regression results

Comparison of original regression results (Table 1) with those of estimation on reduced set: "net price" (Table 3) and actual price

Independent variable: Coefficient significant in all three equations unless noted otherwise	Coefficient from Table 1 (original set)	Coefficient from Table 3 (reduced set)	Coefficient from actual P (reduced set)
Ploughland	333	433	400
Garden	1813	1731	1669
Meadow	808	806	806
Pasture	578	655	588
Woodland	317	307	351
Water**	-71.9	471	1213
Moor***	-1251	-5770	-268
Waste**	-2771	-5884	-5773
Farmyard****	3232	5147	3327
Building value (marks)	54.2	23.6	35.1
Trend*	4.13	3.74	4.05
Property over 100 ha.*	150	153	161
Relative*	-163	–	-183
Constant term****	-1168	230	-6.73
Standard error of estimate	12678	11613	12169
R^2	0.968	0.973	0.971
Number of observations	438	401	401
Degrees of freedom	424	388	387

*See text for explanation of this variable.
**Insignificant in original equation. This coefficient significant in the other two equations at 95% level or better.
*** Insignificant in both regressions for actual price paid (full data set and reduced set), but significant in the regression for "net price." **** Insignificant in all three equations.

Table 5: Final form of the regression equation, full data set

Regression with deflated purchase price as the dependent variable Mean = 40401 Marks, Standard deviation = 70272 Marks, Weighted mean price per hectare = 762 Marks

Independent variable (unit of measurement is hectares unless otherwise noted)	Coefficient	Std. Err
Ploughland	342	49.0
Garden	1775	625
Meadow	822	101
Pasture	580	80.7
Woodland	303	123
Water	-65.8	74.8
Moor	-1527	1053
Waste	-2828	1898
Farmyard	4030	3241
Building value (marks)	49.4	5.88
Trend*	4.29	0.495
Property over 100 ha.*	145	27.9
Relative*	-161	24.3
Dummy for existence of business (1 or 0)*	5051	2724
Constant term / standard error of estimate:	-1270	12732
R^2	0.968	
Number of observations / degrees of freedom:	438	423

* See text for explanation of these variables.

marks per hectare, or about 19% of the mean price. The trend remained significant and positive, although small: at just over four 1913 marks per hectare per year, it was less than 0.6% per year measured at the mean price. The value of buildings as a factor in price determination remained strongly significant across all formulations of the regression equation: in the final form buyers paid just over 49 marks for each mark of assessed use value of buildings.

Generalizing from these results is problematic, however. Angerburg was a rather remote place in East Prussia, with relatively poor land transport[24] and therefore a looser connection to the domestic market than would be the case for a similar riding in Brandenburg or the province of Saxony, for example. The premium paid for large properties could well be a local phenomenon, perhaps even an anachronism, that could exist precisely because of the riding's remoteness. It may have disappeared, or not have existed, in other more central parts of Prussia. Nevertheless, the consistency with a phenomenon such as the formation of exclusive groups which could exclude others on the basis of high prices, as suggested by Morgan (1997), is intriguing, and bears looking into in other parts of Prussia and Germany as a whole.

There seems far less reason to doubt that the »family discount« was - indeed remains today - a general phenomenon. It would not have been on a government form used all over Prussia if it had not been seen to be widespread. What we have done here is to get an estimate of its size: on the order of 20 percent, on average. That may well be close to the mark in other parts of Prussia as well.

That such a simple formulation could explain nearly all of the variation in purchase prices of landed properties is encouraging. The very fact that variables measured in area terms explain so much of the variation in value reinforces the point made in an earlier article that area is a very good proxy for value in the distribution of landed properties (Eddie 1993). That paper was concerned with *relative* values; we have here results which would also allow us to calculate *cardinal* values: While it is beyond the scope of the present paper, these results could be used to estimate land values for whatever area one feels comfortable in generalizing about from the Angerburg results.

Finally, it is my hope that this study might stimulate others to investigate price formation in agriculture, particularly in Prussia. The relationship of the prices of large and small properties is particularly of interest, in that so much of the analysis and the debate about the development of agriculture in 19th-century Prussia focuses on the question of large vs. small properties, in the political as well as in the economic sphere (Achilles 1993, chs. 3 and 4).

References

Achilles 1993: Walter Achilles, *Deutsche Agrargeschichte im Zeitalter der Reformen und der Industrialisierung* (Stuttgart: Ulmer).

Eddie 1993: Scott M. Eddie, »The Distribution of Landed Properties by Value and Area: A Methodological Essay based on Prussian Data, 1886 - 1913,« *Journal of Income Distribution,* vol. 3, no. 1 (Summer 1993), pp. 101 - 140.

[24] See for example the brief description of Kreis Angerburg in the various editions of Ellerholz 1884ff.

Eilerholz 1884ff: *Handbuch des Grundbesitzes im Deutschen Reiche* (Paul Ellerholz, general editor), East Prussia, 2nd, 3rd, and 4th editions, Berlin: Paul Parey, 1884, 1895, 1907.

GStPK Rep. I 151 II: Geheimes Staatsarchiv Preußischer Kulturbesitz, Berlin-Dahlem, Rep. I 151 II, nos. 3728/1, 3728/2, and 3728/3 (Finanzministerium).

Jacobs-Richter 1935: Alfred Jacobs and Hans Richter, *Die Großhandelspreise in Deutschland von 1792 bis 1934,* »Sonderhefte des Instituts für Konjunkturforschung,« Nr. 37 (Berlin: Hanseatische Verlagsanstalt Hamburg).

Lesigang 1900: Lesigang (no first name given), »Grundsteuer,« in *Handwörterbuch für Staatswissenschaft,* 2nd. ed., vol. 4 (Jena: Gustav Fischer).

Morgan 1997: Peter B. Morgan, »Two-sided search and matching,« unpub-lished paper.

PSL 1917: Preußisches Statistisches Landesamt, »Die Kaufpreise der ländlichen Besitzungen im Königreich Preußen von 1895 bis 1912, unter besonderer Berücksichtigung der Jahre 1910 bis 1912,« *Zeitschrift des Königlich Preußischen Statistischen Landesamts,* Ergänzunsheft 44 (Berlin: 1917).

Sering 1910: Max Sering, *Die Verteilung des Grundbesitzes und die Abwanderung vom Lande* (Berlin: Paul Parey).