CEC: Modeling the Residual Market Value of Construction Equipment under Changed Economic Conditions

Gunnar Lucko, A.M.ASCE

*Catholic University of America*

**Abstract**

This paper reapplies a statistical model that was created before 2003 to forecast the residual value of used heavy construction equipment in the U.S. The objective is to evaluate the performance of the model in the radically changed economy in the second half of the decade. The research hypothesis addresses whether the model is still usable and functional even almost a decade after the data occurred that determined its coefficients. This existing statistical model was a comprehensive multiple linear regression analysis for various categories of common types and sizes of equipment. Manufacturer, condition rating, and auction region are included as binary indicator variables. Statistically significant macroeconomic indicators can be used directly, but provide new challenges as various governmental data series have been discontinued or modified, so that valid replacements must be found that range until the time of the economic crisis.

Performing such a re-analysis faces technical challenges, including that several macroeconomic data series since the time of the original study have been discontinued or modified by their governmental sources. Moreover, new auction sales records mostly lack conditions and are devoid of locations. Various reconstructions are explored to still enable a valid re-analysis and upon validation are used to augment the previous data series seamlessly with new data. Expert elicitation is used to select several equipment types that are likely to be affected by the recent economic downturn. The existing implementation tool is provided with the updated data for calculating forecasted residual values that are compared statistically with actual auction sales prices. Both practical and numerical problems are identified. The model
consistently underestimates the actual values, indicating that they are less affected by the economic crisis or new variability is introduced from reconstructing its inputs. It is recommended that such model should be regularly updated, ideally by a professional organization.

**Keywords:** Construction equipment, cost analysis, regression analysis, data analysis, economic factors

1 Introduction

Construction equipment managers must make sound decisions on how, when, and under which conditions to acquire, operate, and dispose of machines. Valuation, the accurate forecasting of the so-called residual value of used heavy construction equipment without actually selling it in the market, poses a tremendous challenge for the construction industry and its construction engineering and management specialty. Both owners of equipment fleets and users of rental equipment are impacted financially by the residual value. For owners, the cash inflow from a possible sale determines the *economic life* of the machine. It is reached at the point in time when the composite cumulative average of owning and operating costs has a minimum value or maximum profit. Beyond this time it is no longer economically feasible to own a unit of used equipment. The residual value thus directly influences the timing of a business decision to sell.

An earlier major study addressed the lacking availability of a statistically sound yet user-friendly forecasting model for the residual value of different types and sizes of equipment (Lucko 2003) and is summarized in the literature review of the following section. Since then, a small number of additional studies either focused its applicability specifically to older machines (Burstein 2005), hedonically analyzed the price components of machines and linked them with the revenues from auction sales (Ponnaluru 2009), or used a datamining approach to extract a model from a large data sets (Fan et al. 2008). The recent severe economic crisis and subsequent recession, however, open a set of new questions regarding the performance of the existing statistical model that had been created by this author. They are related to the continuity, applicability, and performance of the model, which is the focus of this paper.
2 Literature Review

2.1 Importance of residual value

The residual value plays an important role in the total hourly costs that the owners charge internally to the account of the machine. If the owner self-performs operations with the machine, the amount of income that the machine needs to generate for the owner to break even is equal to the difference of the actual purchase price – usually discounted from the published list price of the manufacturer – minus the residual value and plus any desired profit margin. If the machine is rented externally, the user, for example an earthmoving contractor, will have to pay these total hourly costs before earning any profit from working on a heavy construction project. The minimum value of the hourly rate that said user can charge to break even for performing their contracted services therefore is also completely determined by the residual value. The only financial difference between the former and the latter scenario is that the total hourly costs will have to include the profit margin of the owner plus the additional profit margin of the user who rents the machine to generate a revenue stream.

The residual value therefore is of fundamental importance for any operation that involves heavy equipment, regardless of whether it is owned or rented, and it is imperative for the economic survival and long-term success of any contracting companies that use heavy equipment to have at their disposal a reliable forecasting tool, which is supported by carefully maintaining records of their equipment-related costs. This paper focuses on analyzing the long-term performance and practicability of such a model.

Beyond its aforementioned uses in economic cost calculations, the very nature of the residual value itself contributes further to its importance. Among the various components of owning costs, including sales fees, financing interest, insurance premiums, and property taxes, the residual value has been the most elusive and difficult to predict element (Lucko et al. 2007). Its special nature stems from the fact that it is not an actually realized dollar amount from an actual market transaction, which would set the price. It is rather a forecasted value that is based on many independent variables of the machine itself, for example its make, model, and condition. Additional factors are the current market within which it could be sold,
the situation of the overall national and even international economy, the regional and local conditions of
the construction industry, and the specific supply and demand that is created by the interaction of several
contracting companies that may seek to sell or purchase the specific unit of used heavy equipment.

2.2 Previous studies

2.2.1 Related studies in construction equipment valuation
Previous work by Kastens (2002) created a calculator tool for analyzing the owning and operating costs of
heavy construction equipment that used an empirical relationship to forecast the residual value component
of owning cost. Besides this inverse square root function of age measured in hours, it also included
financing interest, insurance premiums, licensing costs, and property taxes, the latter of which were based
on an assumed straight-line depreciation of book value of the machine. It also included forecasted
maintenance and repair costs as per a quadratic function of age that was developed by Mitchell (1998).
The following sections summarize the approach of the original study (Lucko 2003) for this re-analysis.

Additional studies that continued the investigation of valuation for used heavy construction equipment
were performed by Burstein (2005), Fan et al. (2008), and most recently by Ponnaluru (2009). A current
study on life-cycle analysis is noted here, but is ongoing and unpublished at the current time (May 2010).

Specifically, Burstein used the original dataset of Lucko (2003) and selected only those machines over
4 and 5 years of age, respectively, and performed a new regression analysis of the residual values of older
equipment. It found that the original model still performed with the best ‘goodness-of-fit’ and constituted
a feasible forecasting tool for the residual value of used heavy construction equipment. Interestingly, the
overall ‘flatter’ curves with small decreasing slopes in the models that were refitted to the cropped data
sets proved to become more susceptible to variability in the explanatory variables, which in turn reduced
the predictive capability of the refitted models. It was suggested that larger data sets should be collected
and analyzed, and that a more detailed exploration of annual use of different types of equipment is needed.

Taking a completely different approach, Fan et al. (2008) applied a data mining approach on a large
but less structured data set that contained residual values of used heavy construction equipment. Their
approach did not separate the data set into categories by equipment type and size, but rather automated the analysis. While data mining allows gleaning quantitative information from such data sets and through its automation requires only “minimal user input” (Fan et al. 2008), the tradeoff for analyzability is a decreased or even lacking interpretability. Splitting data sets into categories with associated individual explanatory variables by such a data-driven approach may not create intuitively understandable modelling concepts. In fact, “[t]he inference capability and storage mechanism of data mining greatly loosens the problem scope definition compared with the traditional statistical methods” (Fan et al. 2008). To lessen this potential problem an initial distinction by type therefore still had be made manually. Moreover, individual models will structurally differ from each other, which may reduce their usability in practice.

Most recently, an economic study performed three separate analyses, a hedonic price analysis of used excavators[...], an empirical estimation of revenue functions of auction houses[... and an] empirical analysis of price relationships in a multi-item, multi-type auction (Ponnaluru 2009). The first of these analyses used “[s]patial hedonic models [that] have been widely used in real estate economics” (Ponnaluru 2009) and also considered the possible impact of the auctions. Is results confirmed that the manufacturer, condition rating, and auction region significantly impact the actual price as postulated by Lucko et al. (2006), as well as the auction houses and even the individual auction events themselves.

2.2.2 Auction sales of used construction equipment

The major original study upon which this update and re-analysis is based (Lucko 2003) collected large data sets from commercial auctions of 11 different commonly used types of heavy construction equipment, some of which had several subcategories. These types included track and wheel excavators, wheel and track loaders, backhoe loaders, integrated toolcarriers, rigid frame and articulated trucks, track dozers, motor graders, and wheel tractor scrapers. Both the selection of types and makes were deliberately limited in their respective scope to commonly found machines; (a) to ensure a wide applicability of the final model and (b) to ensure the availability of sufficiently large data sets to generate a statistically significant model. Data were thus limited to four leading manufacturers, including Caterpillar, Deere, Komatsu, and
Volvo so that up to four different makes existed for each different type. Only wheel loaders were actually produced by all four manufacturers; the number of makes for the other types ranged from one to three.

The type categories were subdivided further into size categories by using a ‘characteristic measure’ that captures the primary productive function of the particular type, for example transporting or transforming material by hauling or pushing it. The chosen measures were standard operating weight (empty) for track and wheel excavators, rigid frame and articulated trucks, and wheel tractor scrapers, general purpose bucket size for wheel and track loaders and backhoe loaders, and net horsepower (flywheel) for integrated toolcarriers, track dozers, and motor graders. Bucket size was not used for excavators as their attachments are exchanged frequently so that no standard setup existed at auction. Information gleaned from the brief textual descriptions that were provided as part of the auction records confirmed such variations. The number of size categories ranged from one to five and all intervals except had an equal range except for the largest one that had an open upper boundary.

The chosen cutoff values between categories reflected the desire to match industry-common size classifications with data sets of similar magnitude within each subcategory. Overall, the 35,542 data points (reduced to 35,202 after outlier elimination) from auction sales comprised 11 common types with 28 size categories. Table 1 lists those equipment types that are likely to be affected by the current economic crisis as per the expert elicitation that is explained further below, the numbers of data points for the categories, and their normal and adjusted coefficients of determination ($R^2$ and adjusted $R^2$) that measure the ‘goodness of fit’ of the regression model to the data. The $R^2_{adj}$ was created as per Eq. 1 to correct for the number of explanatory variables to prevent over-fitting the regression model (Lucko 2003). Its value is always smaller than $R^2$ and is more realistic for comparing the performance of models that contain different numbers of explanatory variables as was done in the original study during the variable selection mechanisms. Either one can be used in this re-analysis as its models have the same variables.
The auctions for which data were available ranged from January 15, 1994 to September 28, 2002. Only machines of up to and including 15 years of calendar age at the time of their auction sale were included; all other entries were removed from the overall data set. Accordingly, the respective year of manufacture (validated via the unique identifier of the serial number, which are awarded sequentially) predated it and ranged from 1979 to 2002. While most economic data were available with a resolution of monthly values, the fact that no exact data of manufacture was recorded or reconstructable from any incidental data resulted in the final model only being able to have a resolution of one calendar year for its input variable of ‘equipment age’ and the final model can be interpreted as providing a somewhat ‘stepped’ approximation of the actual aging process and its associated decline in the condition rating.

Auction data were obtained from two different sources that continue to be marketed under the names of LastBid® (EquipmentWatch 2010) and TopBid (Randall-Reilly Publishing 2010). “Both data sources gave written permission to use their data for research purposes” (Lucko 2003). Validation performed under the original study confirmed that these two sources were indeed independent, identifying even what appeared to be individual manual transcription errors that differed between them (Lucko and Mitchell 2010). The significant overlap between the data from both sources allowed an internal validation, which corrected a small number of transcription errors and made the affected data still usable. Table 1 lists ‘demographics’ of the data sets without outliers from which coefficients were calculated. A detailed description of how to eliminate outliers is provided in Lucko et al. (2006).

\[ R^2_{adj} = 1 - \left( \frac{SS_{err}}{n-p} \right) \frac{SS_{tot}}{n-1} < R^2 \]  

(1)
2.2.3 Economic conditions during auction sales

Data from auctions in the U.S. and in Canada were included originally; while other nations were assumed to have sufficiently different conditions for equipment auctions to merit their own more detailed studies.

Measuring or modeling the behavior of the overall economy is a challenging task (Gray 2009) and necessarily only provides an incomplete view of such complex a system. For the previous study and its update, “indicators were selected based on their general acceptance as measures of the state of the economy and their applicability for the Construction Industry, their public availability from official sources, and their frequency” (Lucko 2003). These data comprised series published by the U.S. Bureau of the Census, the Federal Reserve Board, the Bureau of Economic Analysis, the Bureau of Labor Statistics, and private sources that are published in regular trade magazines of the construction industry.

The potential macroeconomic data series that could be selected for inclusion in the model covered a broad range of factors, ranging from the gross domestic product (GDP) via interest rates of 10-year maturity securities, industrial production of construction steel, sales of new single family homes, and composite stock price indices to the construction cost index published by Engineering News Record magazine and numerous others (Lucko 2003) and covering weekly, monthly, and quarterly data series.

Several pairs were found to have very high correlations with each other as measured by the Pearson coefficient of correlation $R$. Among them was 0.99665 for gross domestic product (nominal, seasonally adjusted) and total retail sales (seasonally adjusted), 0.99495 for total retail sales (seasonally adjusted) and the consumer price index for all items of urban consumers (1982-84 = 100, seasonally adjusted), and 0.99219 for the consumer price index and the producer price index for finished goods (PPI, 1982 = 100, seasonally adjusted). Regression models should not contain such redundant explanatory variables, which would complicate determining how individual explanatory variables contribute to the overall response. This so-called multicollinearity problem (Montgomery et al. 2001) was avoided by using the so-called forward selection, backward elimination, and stepwise selection mechanisms that are explained further below, to identify which of the potential variables should be included in the final regression model.
Based on an analysis of the goodness-of-fit in relation to the number of macroeconomic indicators, which show an asymptotic behavior, it was decided to include two indicators in the model (Lucko 2003) to provide a broader coverage of the complex behavior of the overall economy than one indicator alone.

2.2.4 Trade models with selected macroeconomic indicators

In order to permit industry practitioners to use the regression model more easily, the initial full model was modified into two different versions; one ‘plain’ version devoid of any macroeconomic indicators to allow a machine-centered prediction of the residual value based only on its age, manufacturer, condition rating, and auction region and one ‘trade’ version that used a shorter list of nine potential macroeconomic indicators including GDP, interest rates, housing starts, and total construction put in place, which were chosen due to their availability as “commonly reported in trade journals, thus reducing the effort of keeping a current economic database for the user. The smaller selection of economic indicators also helps avoiding multicollinearity problems caused by these indicators” (Lucko 2003). Within the tradeoff of ease of applying the model, availability of its inputs, and the ability to still obtain a high goodness-of-fit, the trade models were chosen for this re-analysis. Their coefficients of determination $R^2$ fall “between the values for the two other types of regression models” (Lucko 2003).

The final trade model that was created from the statistical regression analysis had the form of Eq. 2.

$$RVP = \beta_0 + \beta_2 \cdot (age)^2 + \beta_1 \cdot age + M_1 \cdot m_1 + M_2 \cdot m_2 + M_3 \cdot m_3 + C_1 \cdot c_1 + C_2 \cdot c_2 + C_3 \cdot c_3 + R_1 \cdot r_1 + R_2 \cdot r_2 + R_3 \cdot r_3 + E_1 \cdot e_1 + E_2 \cdot e_2$$

(2)

where $RVP$ is the residual value percent that normalizes the residual value by dividing it by the list price, $\beta_0$ through $\beta_2$ are regression coefficients ($\beta_0$ being the intercept), $age$ is in calendar years, $M_i$, $C_i$, and $R_i$ are coefficients for manufacturer, condition rating, and auction region, $m_i$, $c_i$, and $r_i$ are respective binary indicators, $E_i$ are coefficients for economic indicators, and $e_i$ are their economic indicator values.
2.2.5  Statistical analysis and model development

A multiple linear regression analysis was performed separately for each of the 28 size categories for the independent explanatory variables (model inputs included the “age in years, the indicator variables manufacturer, condition rating, and geographic region, and selected macroeconomic indicators”) and dependent response variable of “residual value percent, defined as auction price divided by manufacturers suggested retail price” (Lucko 2003). The reason for normalizing the output toward a residual value percent was to ensure comparability and be independent of inflation. List prices, also called manufacturers suggested retail prices, were obtained from equipment manufacturers and compiled into a catalogue that – along with economic data – was matched with the auction data before model analysis. It continues to be assumed that these list prices (which formally are only ‘values’, as they do not result from actual market transactions) are assigned by the different manufacturers in a sufficiently consistent way. Condition ratings, which serve as proxies for the effects of the wear and tear from age and cumulatively increasing usage, also continue to be assumed to be assessed consistently by appraisers with standard checklists (Lucko 2003). However, anecdotal information from an equipment management professional indicates that there may be significant fluctuations in list prices and discounts off them and complex multiple discount structures to complicate discount-percent-based competitor pricing and deliberately disguise actual initial purchase prices, which may have fluctuated less as they grew consistently. Without any detailed study on the relationship between list prices (or rather advertised values for marketing and discounting purposes) and actual initial purchase prices, within the scope of this re-analysis it is not possible to correct for such potential deviations and thus the aforementioned assumption is maintained.

Among the potential models of different polynomial orders, exponential, and logarithmic functions, a second-order polynomial function of age with additive linear components was chosen to (a) model the anticipated approximately asymptotic decay in residual value that has been observed in real machines, while (b) making the model only as complex as necessary – in the spirit of Ockham’s razor – to ensure its usability by practitioners and be fully transparent. The tradeoff of thus incurring increasing residual
values on the right half of the parabola was handled by limiting all of their ages to 15 years or less and by automatically truncating this quadratic function in the implementation of the model, a spreadsheet tool.

Outliers are defined as those “extreme observations [that] significantly differ from others by their sign or magnitude and are easily identified in a scatterplot of the data”. Since they are “inconsistent with the basic relationship captured by the other data points” and may even act as influential points that can individually distort the model away from the overall behavior (Lucko 2003). Reasons for outliers can be errors in measuring or recording data, or a changed scope, for example in the setup of a machine. A detailed manual review of the data found various outliers from non-standard setups at the time of auction.

They were eliminated based on having excessively large scaled residuals that deviated in their value from the regression model. The total of 340 outliers was only 0.96% of the data in the initial data set, an indication of the consistency of the dataset. Final model coefficients were the recalculated without them.

Forward selection, backward elimination, and stepwise selection were all employed to select the two macroeconomic indicators from the long list of potential ones that were part of the original data set. They function either by testing all potential indicators, adding the one that most increases the $R^2$, and repeating (forward selection), subtracting the one that most decreases the $R^2$ and repeating (backward elimination), or alternating between adding the most significant and then dropping the least significant indicator from the model. Overall, the selections made by these three mechanisms were very consistent (Lucko 2003).

3 Need for Updated Study

Starting in approximately 2006 the U.S. economy has suffered from a massive nationwide crisis that has even spread internationally to numerous other Western countries. It can be described as having started in the U.S. with rapidly increasing numbers of defaults on subprime and adjustable rate mortgages that reset to higher rates as the earlier consistent growth of housing prices and related possibilities of refinancing abated. It followed that numerous borrowers were unable to service even interest payments and incurred negative equity. Despite the inherent risks of such loans these were neither reflected in their ratings nor in
their pricing structure. Highly complex derivative investment products containing such loans were created and traded internationally in large volumes (Gray 2009; Pressacco and Seravalli 2009). Immense losses of global stock values, bankruptcies of numerous small and large U.S. commercial and investment banking institutions, which had been subject to significantly less regulatory oversight, followed. An enormous and controversial government-initiated ‘bail out’ program for banks that had grown to volumes that were deemed ‘too big to fail’ and for U.S. automotive companies became necessary, lowering the interest for federal interbank lending to virtually zero. U.S. currency valuation fluctuated as a major recession began.

It is outside of the scope of this paper to provide a detailed description or an analysis of this economic crisis itself. Despite historically unprecedented government interventions to provide the financial markets with liquidity, government-initiated mortgage intervention programs, and planned legislation to provide stricter regulations of financial institutions, several of its contributing factors still exist to varying degrees. Furthermore, the medium-term and long-term impacts of this economic crisis on financial, legal, political, economic, and social circumstances in the U.S. are currently (May 2010) still interacting and taking shape.

3.1 Economic crisis and recession

The aforementioned economic crisis has led to an extended recession from which the U.S. and large parts of Western national economies are suffering since approximately 2007, its formally recognized start. Its significance at the national level is underlined by an analysis by the Federal Reserve of St. Louis, which found that “[m]ain recession indicators tend to support the claim that this recession could be the most severe in the past 40 years. However, we are still far from another Great Depression” (Gascon 2009).

Figure 1 shows the GDP since 1960 superimposed to the recessions in the U.S. economy, shaded as vertical gray stripes, that have been officially recognized by the private National Bureau of Economic Research (2010): April 1960 to February 1961, December 1969 to November 1970, November 1973 to March 1975, January to July 1980, July 1981 to November 1982, July 1990 to March 1991, March to November 2001, and ongoing since December 2007 to the current time (May 2010, all dates from peak to trough), although some analyses indicate that it may have ended in July 2009. For the purposes of this re-
analysis, the recession is assumed as lasting to the present time. As can be seen in Figure 1, one of the criteria for defining it is that the GDP stagnates (or in real dollars even declines) during such period in time. Overall, the potentially severe impact of the current economic crisis clearly merits this re-analysis.

3.2 Equipment valuation

Two important aspects of equipment valuation must be distinguished. For the purposes of this re-analysis, ‘value’ designates a theoretically generated forecast that ideally is derived from real data, of what would be the expected market price if the equipment were sold at a specific point in time. This differs from the ‘book value’ concept that is used only for tax accounting and uses non-statistical simplified decay models.

‘Price’ on the other hand designates the result of an actual market transaction between two willing and assumed to be equally informed parties (Lucko 2003). In that sense, list prices should better be referred to as list values or manufacturers suggested retail values, as they are unrealized and are set by manufacturers primarily as the baselines for marketing purposes, for example to create suitable discount percentages.

<INSERT FIGURE 1 HERE>

3.3 Research hypotheses

The research hypothesis for this update and re-analysis of the values compared to prices of used heavy construction equipment is phrased to express the long-term performance and practicability of the model.

*It is possible to continue to use the existing regression model with new publically and commercial input data, including values of macroeconomic indicators, in the current economic crisis to calculate forecasted residual values that match the actual auction sales prices in a statistically significant way.*
3.4 Research methodology

Similar to the internal validation procedure that was applied in the original study to test for consistency within the same data set by splitting it randomly for generating regression coefficients (Lucko 2003), the research methodology of this update and re-analysis separates the newly compiled data set into two parts.

The actual auction prices are kept separate and are not used for the forecasting in any way. Only the age, manufacturers, condition rating, auction region (as binary indicator variables), and the two particular macroeconomic indicators for the specific equipment type and size category are used as shown in Figure 2. These values are explanatory variables for the existing regression model, whose coefficients were fine-tuned during the original study. It is assumed that a continued growth within individual values may have occurred and is acceptable for this re-analysis, as this would simply equate to an extrapolation that is performed with the regression model. Generally, it is desirable for a regression model to forecast within the range of values that its original data set comprised, to whose variability it would have been trained.

Once the newly forecasted residual values are available, they are compared statistically with the actual auction sales prices, which are normalized to percent values by dividing them by their respective unique inflation-extrapolated list prices for the particular type, size, make, and model of construction equipment.

<INSERT FIGURE 2 HERE>

4 Data Collection

4.1 Expert elicitation

The scope of this study was set to choose exemplarily several types of heavy construction equipment that would be most affected by the economic crisis that is strongly related to the U.S. housing market. Three experts in heavy construction equipment were contacted to identify the respective types. They were senior
faculty members who had numerous publications specializing in equipment management. They reviewed the complete list of the aforementioned 11 common equipment types and rank-ordered them from most likely to be affected to least likely to be affected. All experts concurred that equipment that is involved in projects of residential construction and small-scale utility construction would be most affected. The types with the three consistently highest ranks are examined here: Backhoe loaders, wheel loaders, and on-highway rigid-frame trucks. Of these types, trucks had not been considered in the original study as they are manufactured by other companies. Note that a small caveats exist for this re-analysis in that the original trade model for small backhoe loaders had a lower $R^2$ (‘goodness of fit’) and a smaller data set as per Table 1. All other values of $R^2$ exceeded 0.65. Table 2 lists the numbers of data points for these categories and their pairs of macroeconomic indicators.

<INSERT TABLE 2 HERE>

$WTR$ is the public construction put it place of water supply facilities (monthly, Bil. 1996 dollars SAAR), $TTLCNST$ is the total construction put in place (monthly, Bil. 1996 dollars, SAAR), $SWR$ is the public construction put it place of sewer systems (monthly, Bil. 1996 dollars SAAR), $EMPLC$ is the employment in construction (Ths., SA), $INTRST$ is the interest rate in percent per annum for 10-year constant maturity securities, $GDP$ is the gross domestic product (quarterly, Bil. Dollars, SAAR, nominal), and $HMSTS$ is the total privately owned housing starts and building permits (monthly, Ths., SAAR). The macroeconomic indicators $HWY$, construction put in place for public highways and streets (monthly, Bil. 1996 dollars, SAAR), and $PPIME$, producer price index for construction machinery and equipment (monthly, 1982 = 100, SA) were available for the trade models but were not selected for those in this re-analysis. The abbreviation SAAR denotes a seasonally adjusted annual rate (SA, Bureau of the Census 2010a) as the 12-fold “of a monthly value [that] is calculated by dividing the monthly value by an adjustment factor to remove seasonal effects. It assumes that… the particular monthly value would be representative throughout the entire year… while the SAAR does not constitute any measurable
value in the actual economy, it allows direct comparisons between monthly, quarterly, and annual values” (Lucko 2003) that are otherwise impossible due to the different frequencies and periodic changes.

4.2 Challenges of new auction sales data

The two previous sources of auction sales data LastBid® and TopBid provided information that their data cover the previous five and three years, respectively, both of which suffices for this re-analysis. Both provide the important user-friendly option of exporting data sets into an electronic spreadsheet format. While a cost reference guide and even a cost calculator tool are available and performance handbooks from manufacturers include tables of average owning and operating costs, their undisclosed analytical or empirical processes and proprietary nature do not allow ascertaining their forecasting quality statistically.

Prices and timeframes for electronic subscriptions to the auction data have changed since the original study. In this update only one source is used, TopBid. Any data with potential errors are eliminated in the outlier elimination procedure. While the original data set covered both U.S. and Canadian auctions, only U.S. values are included in this update as Canada was affected less severely than the U.S. during the current economic crisis. Limiting the re-analysis to only the U.S. better maintains the internal consistency.

It is noted that the new auction data show a significant decrease in quality as compared to the original study. The new data do not contain any geographic identifiers, for example states, cities, or ZIP codes, for auction events, thus rendering the region component of the model dysfunctional. A review of coefficients for the auction region in the trade models shows that they are all smaller than 0.032. This dimensionless number is small compared to the coefficients of the strongest variable, age, which range from 0.6 to over 0.9. It is assumed that such relatively small additive contribution is permanently set to zero. This range of inaccuracy from missing region data is accepted to still enable a valid re-analysis. Only 486 of the 5,708 data points before outlier elimination (8.5%) contain a condition rating, all other entries were blank. A review of coefficients for the condition rating that their numeric range is larger (up to 0.072 for very good and excellent condition of small wheel loaders) than for auction region. It is assumed that all missing
condition ratings are set to a neutral medium of ‘good’. This assumption curtails the potential inaccuracy and enables the re-analysis.

Additionally, three condition ratings are ‘N’, which is surmised to indicate a non-existing value. These data points are removed, along with 94 data points of the initial 5,708 (1.6%) whose year of manufacture is missing. This prohibits calculating age, the strongest explanatory variable in the regression model.

Verbal descriptions of the setup and specifications appear more standardized than in the original study. They are searched for the words ‘no’, ‘not’, ‘poor’, ‘missing’, and ‘inoperable’ to identify any ‘lemons’ and 17 data points (0.3%) are removed due to poor tires. The original data had more varied notes, especially for missing parts. It is unclear whether this was due to the quality of the machines or their auction records. 502 data points (8.8%) of machines that were older than 15 years of calendar age are removed to retain the same scope as the original model. Individual values may not add to 100% as some data points exhibited more than one fallacy. This data-driven outlier elimination leaves a total of 5,099 usable data points that cover 127 models of equipment.

A more practical challenge stems from that limitation that any downloaded electronic file is cropped at 500 data points under the paid subscription. This caused the additional effort of generating 134 individual files with 1 to 445 data points each, some of which had to split larger categories, which were then merged.

All prices are in multiples of $500 as in the original study, which may be price steps used at auctions. They range from December 1, 2007 to July 31, 2009 to cover the entire recession until its estimated end date. Once the data are obtained, list prices for their year of manufacture are matched and added to the spreadsheet from the existing list price catalog that is inflation-extrapolated with the producer price index for finished goods (PPISP3000.US); new individual macroeconomic indicator values as described in the following major section on U.S. governmental data series are added; and data are categorized by size for the different regression models of Tables 1 and 2 and the existing size parameter catalog (Lucko 2003).
4.2.1 Coding non-numerical categorical explanatory variables

One known challenge in data preparation is the conversion of non-numerical categorical variables that are to be used as explanatory variables. As explained by Lucko and Mitchell (2010), categorical variables that use verbal descriptors, (manufacturer, condition, and region), are qualitative in nature and may even be rank-ordered (excellent, very good, good, fair, and poor condition). Since such form makes them unsuitable for directly being included in the regression model, a conversion into numerical variables is required. Note, however, that it would be incorrect to simply convert them into a sequence, for example \{5, 4, 3, 2, 1\} or similar, as this (a) artificially assumes exact values and fixed ratios between them, when categorical variables are in fact more comparable with linguistic variables of fuzzy sets, a concept where membership to a specific category is a matter of degree and an item can even belong to several categories at the same time, provided the total probabilities add to 100%. A valid statistical way to solve this issue is to convert the category labels into a set of binary indicator variables, which avoids the aforementioned problems but comes at the expense of losing any rank-order between the categories (Lucko 2003). The auction regions as per Table 3 group U.S. states as per the official definition of the Bureau of the Census.

Manufacturer names have been disguised by the letters A, B, C, and D as it is not the intention of this study to compare between manufacturers; Lucko et al. (2006) provides further information on that item. Various other required data preparation steps that have to be applied to make the auction sales data in spreadsheet format usable in a statistical analysis are explained in detail by Lucko and Mitchell (2010).

<INSERT TABLE 3 HERE>

4.2.2 Forecasting current list prices with producer price index

Since neither of the two sources of auction data currently provides list prices, it is necessary to inflation-correct all auction sales prices back to constant November 31, 2002 dollar. The producer price index for finished goods is used again, but in its more precise monthly series rather than the average annual value.
of the original study. This has the advantage that the existing catalog of list prices can be directly used again. Another advantage is that no list prices are inflation-extrapolated to the current time (May 2010), which – due to their anomalously reported high variability – would introduce an additional inaccuracy.

4.3 Challenges of U.S. governmental data series

Despite the perceived permanence and continuity of the publications that the U.S. government produces regularly for a plethora of macroeconomic indicators, it is found that surprisingly some massive changes have been made in by the Bureau of the Census to its former C30 family of data series that measured the value of new construction put in place, which, as listed in Table 4, provides half of the macroeconomic indicators that serve as input for the regression model. As posted on the new website for construction statistics “NOTICE: AS of August 1, 2003, the old types of construction classifications and the constant dollar series are no longer published. New monthly detailed types of construction are available at: http://www.census.gov/const/www/c30index.html” (Bureau of the Census 2010b).

In fact, the website http://www.fedstats.gov/cgi-bin/A2Z.cgi?C of the collection of statistics from reportedly over 100 federal agencies contains a broken link under the entry ‘construction’ that may previously have pointed to the former website http://www.census.gov/const/C30/c30curtb.html but has not been updated in seven years. The assumption of continued availability from the U.S. government that provided the reasoning for creating the aforementioned trade models thus cannot be maintained in the long term. Creators and users of future forecasting models that include any macroeconomic indicators should include provisions that make them flexible in terms of which data series can be used as input. On the downside, this may render such statistical regression model cumbersome to use in practice and a sensible tradeoff between their long-term usability and complexity needs to be found.

Further challenges were added by major changes between the 2002 and the 2007 version of the North American Industry Classification System of the Office of Management and Budget of the White House that is used by the Bureau of the Census, which itself in 1997 had replaced the former Standard Industry Classification system (Bureau of the Census 2010c). “With the changes in project classifications, data
now presented are not directly comparable with those data previously published in the regular-format press releases and tables. Direct comparisons can only be made at the total, total private, total state and local, total federal, and total public levels for annual and not seasonally adjusted monthly data” (Bureau of the Census 2010d). Old data series continue to be available (Bureau of the Census 2010e). Another revision is planned to occur in 2012 (Bureau of the Census 2010f). Such a change of definitions of project types impacts the composition of the pool that an indicator measures and causes a permanent distortion.

Interestingly, not all data series that are published by the Bureau of the Census appear to have been compiled by the government. Some construction cost indices are re-published monthly from Engineering News Record Magazine and quarterly from Turner Construction Company (Bureau of the Census 2010g). Since their exact composition and quality is more difficult to ascertain, this re-analysis only uses major government-compiled data series themselves, for which detailed methodology descriptions are available.

Formally, data points that are published by the Bureau of the Census are considered estimates (having been sampled from various businesses to represent the larger economy) where “[a] necessary part of the process of issuing these early data involves the issuance of subsequent revisions. (...) For total construction, the range of the difference… was from -0.5 percent to 1.6 percent with a median of 0.2 percent” (Bureau of the Census 2010g). For the purposes of this re-analysis, the most recent available revised estimates at the current time (May 2010) were consistently used for all of the data series.

Besides the technical challenges arising from the fact that half of the required input appears to be unavailable for this re-analysis, a more practical challenge stems from the manner in which data are made available to the public in text releases and electronic file formats. Various databases provide full free access (but for example the website Economagic disguises governmental data by adding random artificial digits in white font to prevent direct copying without a paid subscription), but there is no central source for all required major data series. Data for this re-analysis were obtained from the free websites Freelunch (Moody’s Analytics 2010) and Federal Reserve Economic Data (Federal Reserve Bank of St. Louis 2010).
A mnemonic, i.e. a verbal identifier, uniquely labels official data series (Bureau of the Census 2010h) to help verifying that the correct data series is obtained from among the numerous published ones (e.g. in nominal or real dollars, in current or specific-year-based dollars, SAAR, SA, or not-seasonally adjusted).

Table 4 lists the various macroeconomic indicators that are used in this re-analysis, provides details about their past description and source, lists challenges that they present by having been discontinued or modified by their governmental source, and summarizes solutions that are applied to still enable this re-analysis in a statistically valid way by reconstructing and rebasing the respective data series in two steps.

<INSERT TABLE 4 HERE>

4.3.1 Forecasting construction index with inflation index

The monthly data series WTR (public construction put in place of water supply facilities), SWR (public construction put in place of sewer systems), and TTLCNST (total construction put in place) presented particular challenges. Unfortunately, the index that had been used (Bureau of the Census 2010i) to adjust them to billions of 1996 dollars (SAAR) as well as all series in 1996 dollars were discontinued in May 2003. New data series are in current dollars so that any ‘reconstructed’ usable series need to be adjusted also to this constant dollar value (1996 = 100). Fortunately, it is found that the Pearson coefficient of correlation between this “monthly U.S. Census Bureau composite fixed-weighted price index: 1964 to 2003 (1996 = 100)” (Bureau of the Census 2010g) and the producer price index for finished goods for the overlapping range from January 1964 to May 2003 is $R^2 = 0.9722$. There is an excellent goodness-of-fit that allows reconstructing the index in this first step as per Eq. 3. Figure 3 shows their correlation graphically from March 1964 to March 2003 for sorted data pairs. Phases of high variability occurred approximately from September 1985 to August 1987, August 1993 to October 1994, November 1997 to July 1999, and since January 2002. It merits further study, as a relationship with recessions is not evident.
y = 0.8618 \cdot x - 12.098 \quad (3)

where \( y \) is the required ‘monthly U.S. Census Bureau composite fixed-weighted price index: 1964 to 2003 (1996 = 100)’ and \( x \) is the producer price index for finished goods (1982 = 100, seasonally adjusted).

4.3.2 Rebasing construction data series with the reconstructed construction index

Once the required index is reconstructed, in this second step all values for the three data series that are measured in current dollars are rebased to constant 1996 dollars when the index was set at 100. Note that such forecasted values of a series that has been changed significantly can only be a temporary proxy.

However, the reconstructed and rebased three data series are found to not match seamlessly, but while being within the same order of magnitude, still show a ‘step’ after August 2002 when the old data series ended. It is surmised that this is due to a changed pool and assumed that the new proportion is constant, which is justified by the description of how the numerical gap is bridged by using a constant adjustment factor (Bureau of the Census 2010j). They are calculated as 1.8220 for \( WTR \), 1.7891 for \( SWR \), and 4.7066 for \( TTLCNST \) as the averages of monthly ratios for the overlapping range from January to August 2002.

5 Numeric Analysis

Figures 4a through 4e show the pairwise comparisons of the actual residual value percent on the \( x \)-axis versus the forecasted residual value percent on the \( y \)-axis. An accurate forecast is on the diagonal line. Comparisons of the maximum and minimum values, arithmetic means, and standard deviations of these pairs and their differences are given by Table 5 and it is found that the trade models for these equipment types and sizes consistently and systematically underestimate the actual auction values in residual value percent for all sizes of wheel loaders. Worse, numerous forecasted values for the small wheel loaders are
less than zero, which indicates that the carefully applied reconstruction steps for several discontinued or modified macroeconomic indicators may have inadvertently introduced some new systematic inaccuracy.

A hypothesis test is performed on the pairwise matched data in the five size categories. The statistical null hypothesis of this one-tailed t-test is that the arithmetic mean of the differences of forecasted residual values minus actual auction sales prices (both normalized as percent) is zero. The alternative hypothesis assumes it to be negative. In other words, it is tested if the forecast consistently underestimates the actual values. These values are -1.26%, -8.34%, -14.79%, -6.01%, and -10.44% with standard deviations of 9.02%, 5.13%, 5.97%, 6.09%, and 5.87% as per Table 5. Sample sizes as per Table 2 are 3,306, 65, 1,046, 411, and 274 data points, respectively. At the $\alpha = 5\%$ level of significance, the null hypothesis is rejected for all five size categories. It is concluded that the arithmetic means are significantly different, even for the higher variability of the cluster of data points in Figure 4a that appears to be located near the diagonal.

6 Conclusions and Recommendations

The update and re-analysis of this paper illustrates the difficulty of using the established macroeconomic indicators to make valuation forecasts. From the hypothesis test it can be concluded that the model indeed consistently underestimates the actual values. The user-friendliness of the trade models has decreased due to the various challenges that the discontinued or modified data series now present until a usable model is reached. The reliability of reconstructing macroeconomic indicators that no longer exist is doubtful. Other economic models that are based on these types of data series may incur similar long-term problems. For use in practice it would be better to repeat the comprehensive data collection and analysis of the original
study, use only currently available macroeconomic indicators as explanatory variables, and calculate all new regression coefficients. In the long term, it is advisable to institutionalize such statistical analysis. Publishers that collect auction data could automatically analyze their large data sets statistically to update coefficients as an added value for their subscribers, or professional organizations might perform this role. Further research is necessary to determine the timeframe within which a model of this type can provide acceptable forecasts, but beyond that the coefficients should be re-calculated regularly from actual data.

The absence of condition and region data weakens such model. However, both are less important predictors than age. It is recommended to perform research into the exact relationship between age and condition rating, provided a sufficient number of condition data could be obtained. Anecdotal information indicates that conditions may exhibit marketing-psychological bias toward ratings higher than medium and that auctions are becoming global sales events with electronic bidding, so that condition and region indicators should be examined more closely for their contribution to the model and otherwise be omitted.

Quantifying the impact of the various potential new source of inaccuracy is beyond the scope of this re-analysis, however, a number of recommendations for further study can be made, including examining a new set of macroeconomic indicators for their predictive strength and potential permanence, evaluating if an actual decay model may be better suited under the changed economic circumstances, to what degree a non-standard setup is reported and may influence the auction price, whether stratification by manufacturer creates more homogeneous data sets, and finally – and most difficult – to enlighten the relationship between initial purchase prices and list prices, where the latter serve as a proxy of actual but confidential values. Overall, this re-analysis has revealed various practical limitations for a comprehensive statistical model that in the long term in an extreme scenario exceeds the range of data upon which it was trained.
Acknowledgements

The author would like to thank Dr. Michael C. Vorster (retired) of Virginia Polytechnic Institute and State University, Dr. Clifford J. Schexnayder of Arizona State University, and Dr. Aviad Shapira of Technion in Israel for discussions about the usage of different types of construction equipment, Dr. Christine M. Anderson-Cook of Los Alamos National Laboratory for information about regression analysis, Mr. Patrick T. Crail for discussions about industry practices regarding pricing structures, and Mr. Raymond T. Merryman of the Bureau of the Census for helpful information about discontinued and modified data series. The author would also like to thank Ms. Kimberly M. Hoffman, coordinator of science libraries at Catholic University of America, for her assistance in obtaining literature on the current economic crisis.

Notation

The following symbols and abbreviations are used in this paper:

age = age in calendar years
BEA = Bureau of Economic Analysis, U.S. Department of Commerce
Bil. = billion
C = regression coefficient for condition rating indicator variable
c = condition rating indicator variable
cy = cubic yards
E = regression coefficient for economic indicators
e = economic indicator
FRB = Federal Reserve Board
hp = horse power
M = regression coefficient for manufacturer indicator variable
m = manufacturer indicator variable

m³ = cubic meters

N/A = not applicable

n = number of complete observations

R = Pearson coefficient of correlation of the sample

R = regression coefficient for auction region indicator variable

R² = coefficient of determination of the sample

RVP = residual value percent

r = auction region indicator variable

r = studentized residual

SS = sum of squares

SA = seasonally adjusted

SAAR = seasonally adjusted annual rate

Ths. = thousands

x = explanatory (independent) variable

y = response (dependent) variable

α = statistical level of significance

The following subscripts are used in this paper:

adj = adjusted

err = error

t = trade model

tot = total

1 = coefficient for linear term

2 = coefficient for quadratic term
References


Table 1, Distribution of original data from January 15, 1994 to September 28, 2002 up to 15 years old of types likely to be affected by economic crisis after outlier elimination and coefficients of determination in trade models.

<table>
<thead>
<tr>
<th>Type</th>
<th>Bucket Size</th>
<th>Manufacturer*</th>
<th>Total</th>
<th>$R^2$ **</th>
<th>$R_{adj}^2$ **</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[m$^3$]</td>
<td>Caterpillar</td>
<td>Deere</td>
<td>Komatsu</td>
<td>Volvo</td>
</tr>
<tr>
<td>Backhoe loader</td>
<td>0-0.69 (0-0.9 cy)</td>
<td>0</td>
<td>226</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.76 (1 cy) and over</td>
<td>176</td>
<td>7.3</td>
<td>43</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheel loader</td>
<td>0-1.45 (0-1.9 cy)</td>
<td>68</td>
<td>238</td>
<td>131</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>1.53-2.98 (2-3.9 cy)</td>
<td>233</td>
<td>2.1</td>
<td>996</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>95</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3.06-4.51 (4-5.9 cy)</td>
<td>364</td>
<td>104</td>
<td>1,009</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>4.59 (6 cy) and over</td>
<td>210</td>
<td>0</td>
<td>142</td>
<td>88</td>
</tr>
<tr>
<td>Subtotal</td>
<td>N/A</td>
<td>1,051</td>
<td>10</td>
<td>2,321</td>
<td>79</td>
</tr>
</tbody>
</table>

* Note that not every manufacturer produces each type or size of equipment.

** Note that comparatively small values are marked in italics. Values are for trade models.
Table 2, Distribution of new data from December 1, 2007 to July 31, 2009 up to 15 years old of types likely to be affected by economic crisis after outlier elimination.

<table>
<thead>
<tr>
<th>Type</th>
<th>Size</th>
<th>Macroeconomic</th>
<th>Manufacturer*</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[m³]</td>
<td>Indicators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backhoe loader</td>
<td>0-0.69 (0-0.9 cy)</td>
<td>WTR TTLCNST</td>
<td>0 0 0 0 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.76 (1 cy) and over</td>
<td>SWR EMPLC</td>
<td>1,075 2,186 42 0 3,303</td>
<td></td>
</tr>
<tr>
<td>Wheel loader</td>
<td>0-1.45 (0-1.9 cy)</td>
<td>INTRST GDP</td>
<td>22 34 7 2 65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.53-2.98 (2-3.9 cy)</td>
<td>HMSTS GDP</td>
<td>338 462 80 166 1,046</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.06-4.51 (4-5.9 cy)</td>
<td>INTRST GDP</td>
<td>299 19 39 54 411</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.59 (6 cy) and over</td>
<td>WTR GDP</td>
<td>217 0 6 51 274</td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td>N/A N/A N/A</td>
<td>1,951 2,701 174 273 5,099</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Note that not every manufacturer produces each type or size of equipment.
Table 3, Coding categorical explanatory variables with binary indicator variables (adapted from Lucko et al. 2006).

<table>
<thead>
<tr>
<th>Binary</th>
<th>Manufacturer</th>
<th>Condition</th>
<th>Region</th>
<th>States within Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 1</td>
<td>A</td>
<td>Poor</td>
<td>Northeast</td>
<td>CT, MA, ME, NH, NJ, NY, PA, RI, VT</td>
</tr>
<tr>
<td>0 1 0</td>
<td>B</td>
<td>Fair</td>
<td>South</td>
<td>AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV</td>
</tr>
<tr>
<td>0 1 1</td>
<td>C</td>
<td>Good</td>
<td>Midwest</td>
<td>IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI</td>
</tr>
<tr>
<td>1 0 0</td>
<td>D</td>
<td>Very good</td>
<td>West</td>
<td>AK, AZ, CA, CO, HI, ID, MT, NM, NV, OR, UT, WA, WY</td>
</tr>
<tr>
<td>1 0 1</td>
<td>N/A</td>
<td>Excellent</td>
<td>Canada</td>
<td>Not used</td>
</tr>
</tbody>
</table>

Table 3
Table 4, Challenges of current macroeconomic data series used in trade models for selected equipment types and solutions.

<table>
<thead>
<tr>
<th>Label, Mnemonic</th>
<th>Name</th>
<th>Frequency</th>
<th>Source</th>
<th>Unit</th>
<th>Challenge</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTR (CPUWA$96.US)</td>
<td>Construction Put in Place (C30) - Table 5b: Public - Water supply facilities</td>
<td>monthly</td>
<td>Census</td>
<td>Bil. 96$, SAAR</td>
<td>Modified and in current dollars</td>
<td>Reconstruct and rebase to 1996</td>
</tr>
<tr>
<td>SWR (CPUWE$96.US)</td>
<td>Construction Put in Place (C30) - Table 5b: Public - Sewer systems</td>
<td>monthly</td>
<td>Census</td>
<td>Bil. 96$, SAAR</td>
<td>Modified and in current dollars</td>
<td>Reconstruct and rebase to 1996</td>
</tr>
<tr>
<td>TTLCNST (CPU$96.US)</td>
<td>Construction Put in Place (C30) - Table 5b: Total</td>
<td>monthly</td>
<td>Census</td>
<td>Bil. 96$, SAAR</td>
<td>Modified and in current dollars</td>
<td>Reconstruct and rebase to 1996</td>
</tr>
<tr>
<td>INTRST (RGT10Y.US)</td>
<td>Interest Rates (H15): 10-Year Constant Maturity Securities</td>
<td>monthly</td>
<td>FRB</td>
<td>% p.a.</td>
<td>Exact, no issues</td>
<td>N/A</td>
</tr>
<tr>
<td>HMSTS (HST.US)</td>
<td>Housing Starts and Building Permits (C20): Housing Starts: Total privately owned</td>
<td>monthly</td>
<td>Census</td>
<td>Ths., SAAR</td>
<td>Small differences from revisions since 2001</td>
<td>Use most current data</td>
</tr>
<tr>
<td>EMPLC (E23.US)</td>
<td>Form 790: Employment: Construction</td>
<td>monthly</td>
<td>BLS</td>
<td>Ths., SA</td>
<td>Small differences from revisions</td>
<td>Use most current data</td>
</tr>
<tr>
<td>GDP (GDP.US)</td>
<td>Table 1.9 Line 1: NIPA: Gross domestic product</td>
<td>quarterly</td>
<td>BEA</td>
<td>Bil. $, SAAR</td>
<td>Small differences from revisions</td>
<td>Use most current data</td>
</tr>
</tbody>
</table>

* FRB = Federal Reserve Board, BLS = Bureau of Labor Statistics, BEA = Bureau of Economic Analysis
Table 5, Comparisons of maxima, minima, arithmetic means, and standard deviations of actual (A) auction values and forecasted model (M) values and their differences (D) of model minus actual.

<table>
<thead>
<tr>
<th>Type</th>
<th>Bucket Size</th>
<th>Maximum</th>
<th>Minimum*</th>
<th>Arithmetic Mean*</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[m³]</td>
<td>[%]</td>
<td>[%]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backhoe loader</td>
<td>0.76 (1 cy) and over</td>
<td>A 84.12</td>
<td>A 6.27</td>
<td>A 33.58</td>
<td>A 13.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M 55.98</td>
<td>M 8.71</td>
<td>M 32.32</td>
<td>M 8.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D 24.97</td>
<td>D -42.30</td>
<td>D <strong>-1.26</strong></td>
<td>D 9.02</td>
</tr>
<tr>
<td>Wheel loader</td>
<td>0-1.45 (0-1.9 cy)</td>
<td>A 41.75</td>
<td>A 7.51</td>
<td>A 19.76</td>
<td>A 7.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M 34.38</td>
<td>M -6.06</td>
<td>M 11.42</td>
<td>M 7.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D 6.13</td>
<td>D 19.39</td>
<td>D <strong>-8.34</strong></td>
<td>D 5.13</td>
</tr>
<tr>
<td>1.53-2.98 (2-3.9 cy)</td>
<td>A 60.31</td>
<td>A 5.10</td>
<td>A 22.80</td>
<td>A 8.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M 43.03</td>
<td>M -9.89</td>
<td>M 8.01</td>
<td>M 9.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D 6.24</td>
<td>D -35.50</td>
<td>D <strong>-14.79</strong></td>
<td>D 5.97</td>
</tr>
<tr>
<td>3.06-4.51 (4-5.9 cy)</td>
<td>A 58.19</td>
<td>A 5.03</td>
<td>A 23.64</td>
<td>A 8.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M 44.21</td>
<td>M 0.78</td>
<td>M 17.63</td>
<td>M 9.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D 15.66</td>
<td>D -33.93</td>
<td>D <strong>-6.01</strong></td>
<td>D 6.09</td>
</tr>
<tr>
<td>4.59 (6 cy) and over</td>
<td>A 61.56</td>
<td>A 2.54</td>
<td>A 17.69</td>
<td>A 7.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M 30.19</td>
<td>M -8.48</td>
<td>M 7.25</td>
<td>M 6.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D 12.10</td>
<td>D -37.53</td>
<td>D <strong>-10.44</strong></td>
<td>D 5.87</td>
</tr>
</tbody>
</table>

* Note that negative values are marked in italics and underestimated values are marked in bold face.
Figure 1, Gross domestic product (billion dollars, nominal, seasonally adjusted) with superimposed periods of recession since 1960.
Figure 2, Procedure for re-analysis of used heavy equipment values and prices.
Figure 3, Correlation between producer price index for finished goods (horizontal) and composite fixed-weighted price index (vertical).
Figure 4a, Forecasted model values over actual auction values for backhoe loaders 0.76 m³ (1 cy) and over.

Figure 4b, Forecasted model values over actual auction values for wheel loaders 0-1.45 m³ (0-1.9 cy).

Figure 4c, Forecasted model values over actual auction values for wheel loaders 1.53-2.98 m³ (2-3.9 cy).

Figure 4d, Forecasted model values over actual auction values for wheel loaders 3.06-4.51 m³ (4-5.9 cy).

Figure 4e, Forecasted model values over actual auction values for wheel loaders 4.59 m³ (6 cy) and over.
List of Figures

Figure 1, Gross domestic product (nominal, seasonally adjusted) with superimposed periods of recession since 1960.

Figure 2, Procedure for re-analysis of used heavy equipment values and prices.

Figure 3, Correlation between producer price index for finished goods (horizontal) and composite fixed-weighted price index (vertical).

Figure 4a, Forecasted model values over actual auction values for backhoe loaders 0.76 m$^3$ (1 cy) and over.

Figure 4b, Forecasted model values over actual auction values for wheel loaders 0-1.45 m$^3$ (0-1.9 cy).

Figure 4c, Forecasted model values over actual auction values for wheel loaders 1.53-2.98 m$^3$ (2-3.9 cy).

Figure 4d, Forecasted model values over actual auction values for wheel loaders 3.06-4.51 m$^3$ (4-5.9 cy).

Figure 4e, Forecasted model values over actual auction values for wheel loaders 4.59 m$^3$ (6 cy) and over.
List of Tables

Table 1, Distribution of original data from January 15, 1994 to September 28, 2002 up to 15 years old of types likely to be affected by economic crisis after outlier elimination and coefficients of determination in trade models.

Table 2, Distribution of new data from December 1, 2007 to July 31, 2009 up to 15 years old of types likely to be affected by economic crisis after outlier elimination.

Table 3, Coding categorical explanatory variables with binary indicator variables (adapted from Lucko et al. 2006).

Table 4, Challenges of current macroeconomic data series used in trade models for selected equipment types and solutions.

Table 5, Comparisons of maxima, minima, arithmetic means, and standard deviations of actual (A) auction values and forecasted model (M) values and their differences (D) of model minus actual.