A Hedonic Model of Player Wage Determination from the Indian Premier League Auction: Further Evidence[#]

Liam J. A. Lenten, Wayne Geerling[*] and László Kónya

School of Economics and Finance

La Trobe University

Abstract

A range of cross-sectional models are estimated with a view to establishing the factors that determine the valuation of professional athletes in a highly-specialised sport, with an application to cricket's Indian Premier League (IPL). We distinguish between personal characteristic and playing ability factors, and with respect to the former, between ability in different forms of the sport. We find a number of interpretable variables that have explanatory power over auction values, while decomposition according to batting and bowling specialisations produces very different results depending on the use of either Test or One-Day International (ODI) variables. There is also possible evidence of inefficient bidding, insomuch that overbidding was somewhat correlated with players with higher realised values.

JEL Classifications: C21, D44, L83

Keywords: Demand for Sport, Cricket, Auctions, Cross-section Models, Labour Demand

---

[#]Earlier versions of this paper were presented at: (i) Economics Society of Australia (Tasmanian Branch), Hobart, 18 March 2009; (ii) Staff Developmental Workshop, School of Economics and Finance, La Trobe University, 23 April 2009; (iii) Seminar, University of Canterbury, Department of Economics, New Zealand, 1 May 2009; (iv) Seminar, University of Queensland, School of Economics, 22 May 2009; and (v) 38th Australian Conference of Economists, University of Adelaide, 28-30 September 2009. The authors would like to thank the various participants of the seminars and workshops for their comments and suggestions, as well as Seamus Hogan, Michael Smiddy, Ray Stefani, and two anonymous reviewers.

[*]Contact details: School of Economics and Finance, La Trobe University, Victoria, 3086, AUSTRALIA. Tel: + 61 3 9479 1148, Fax: + 61 3 9479 1654. E-mail: w.geerling@latrobe.edu.au

1. Introduction

The inaugural Indian Premier League (IPL) Season of 2008 marked arguably the biggest business revolution in the sport of cricket in the 130-year formal history of the game. As a new start-up league, several design features had to be established. One such feature was the method of allocating players to teams. The IPL decided to conduct a player auction, held on 20 February 2008 in Mumbai, whereby the winning bid would essentially represent the player's wage for the tournament. Such an auction represents an extremely rare opportunity to measure the true value (marginal revenue product) of labour of professional athletes in the sports industry.

Following this opportunity, the purpose of this study is to estimate a hedonic model of player wages arising from this one-off auction process.[1] We wish to use player career statistics at the time of the auction to identify a set of playing traits that are valued most highly by teams. However, we also wish to consider other observable personal characteristics that may have some material effect on players' valuations – characteristics that determine a player's 'marketability' distinctive from their playing talent. While Karnik (2010) performs a similar exercise via extreme bounds analysis (EBA), we use an alternative methodology that we believe to produce noticeably superior fit, as well as addressing the diagnostics of the model explicitly. As will be shown, our results reveal a most fascinating story about the bidding teams and the cricketers involved in the auction.

[1] It is important to note that while further auctions would be planned for future seasons, such auctions would largely involve only new or 'uncapped' players, while there would also be a transfer 'window' (like those in European football leagues), in which existing players could transfer from one IPL franchise to another prior to the commencement of the new season. There was a 3-year lock-in period for players auctioned in 2008 and their contracts were guaranteed by the BCCI.

The structure of this paper proceeds in the following manner. Section 2 summarises some previous literature. Section 3 provides a detailed description of the rules and restrictions of the auction, making possible some inferences based on what auction theory would say about the outcomes involved. Section 4 outlines the data and the econometric modelling used, which leads to section 5, where the results are presented and discussed in detail. Section 6 concludes on a general note.

2. Previous Literature

The IPL formation was an attempt to capitalise on the explosive growth in demand arising from the newer, shorter form of the game ('Twenty20') in recent years at the expense of Test Matches (see Lenten, 2008), as can be seen in table 1 at international level and even more so in table 2 at domestic level. It was also a response by the national governing body, the Board of Control for Cricket in India (BCCI), to mitigate the threat posed by the rival and non-sanctioned Indian Cricket League (ICL). As a measure of the financial optimism surrounding the IPL, the eight franchises were auctioned in January 2008 for a combined total of USD724 million, while the 10-year broadcast rights were sold for USD1,026 million (plus regional rights). The timing of the player auction itself helped create extra hype in advance of the first game.

An auction setting is very useful for reflecting true valuations of cricketers because none of the usual rigidities typically present in the players' labour market in ongoing leagues were present in this case. Refer to Rosen and Sanderson (2001) or Downward, Dawson and Dejonghe (2009) for a description of these restrictions, while the latter provides a textbook coverage on econometric findings in the labour market (see pp. 315-323). Measuring true marginal revenue product is even more

problematic in individualistic sports due to the commonality of rank-order tournaments, and the associated fixed prize money allocation rules. See Leeds (1988) or Ehenberg and Bognanno (1990) for an outline of some of the associated issues, or Frick (2003) for a literature survey of contest theory in sport. However, it is worth noting that the sports economics literature is rich with contributions that utilise player wage data, but usually with a specific application, see Khan (2000) for a general overview of these studies. More specific examples include testing the link between wage expenditure and (on-field) success, such as Berri, Schmidt and Brook (2006), or whether wage discrimination exists in certain leagues (see Bodvarsson and Brastow, 1998; Hill, 2004; or the recent evidence of Goddard and Wilson, 2009).

While the current paper has some conceptual commonality with the *Moneyball*-style setting of Lewis (2003), the use of hedonic modelling to identify determinants of auction values in sport is not uncommon. See, for example, Parsons and Smith (2007), who undertake such an exercise for thoroughbred yearling sales in horse racing. However, such markets are typically largely unrestricted, unlike the IPL auction case. Therefore, we also wish to consider the role of restrictions on the auction in terms of limits on squad composition and overall player expenditure, in an attempt to observe how player valuations are distorted by such restrictions. While these restrictions form a crucial point of distinction of the IPL auction setting, an important survey worth consulting on the more standard labour-market settings in the far less-restricted European football leagues is Frick (2007), which has a short commentary on player salaries by position, analogous to specialisations in cricket.

Despite these related studies, the quantity of other academic work within the sports economics field focussing on the distinctive nuances of international cricket is minimal, a deficiency that this paper hopes to remedy. A general commentary on the economics of cricket can be found in Preston (2006). However, other studies, such as those by Brooks, Faff and Sokulsky (2002); Bhattacharya and Smyth (2003); and Allsop and Clarke (2004), tend to be quite specific in nature, insofar that they each focus on addressing their own distinct cricket-related problem.

3. Background on the 2008 IPL Auction

Eight franchises took part in the original IPL auction, each of which represented a city: Bangalore, Chennai, Delhi, Hyderabad, Jaipur, Kolkata, Mohali and Mumbai. The franchise owners bid on a total of 77 players who had been contracted to the BCCI at the auction that was held on 20 February 2008.[2] There were two exceptions to the bidding process:

i) The IPL placed 'icon' status on five marquee Indian players, not up for auction who were to represent the city where they are from. These players were to be paid 15 per cent more than the next highest paid player in their respective franchises, creating affiliated values in the auction process

ii) Each franchise could only select a maximum of two Australian players (owing to a clause in a contract with Cricket Australia).

Prior to the auction, players were divided into six categories based on their skills, with the 'marquee' players auctioned in the first round. The categories were: opening

[2] To be eligible for a BCCI contract, foreign players required a 'No Objection Certificate' from their country's governing body. Since the 18 April-1 June IPL season clashed with the English County Championship season (as well as a tour by New Zealand), the England and Wales Cricket Board (ECB) refused to release their players. Subsequently, there were no English players signed in the first auction, although non-international regular Dimitri Mascarenhas was bought in the second auction.

batsmen, middle-order batsmen, pace bowlers, spin bowlers, all-rounders and wicketkeepers. The nature of the auction was sequential - chits with players' names were selected from a bowl and franchisees given the opportunity to bid for them. If no-one bid for a player at reserve initially, that chit would be kept separately (the player relegated to a reserve pool to be auctioned later) and the bidding for the other players continued.

The auction itself was carried out over eight rounds with an additional round for players in the reserve pool, making the auction multi-stage. Seventy-five players received bids: Mohammad Yousuf (Pakistan) and Ashwell Prince (South Africa) were the only overseas players not to receive bids at reserve (hence missing out on playing altogether). A total of USD36.78 million was spent (at an average of just under USD500,000 per player). Including the premiums for the 'icon' players, the total wage bill amounted to nearly USD42 million. By country of origin: India 25, Australia 13, Sri Lanka 11, South Africa 10, Pakistan 7, New Zealand 5, West Indies 3, Zimbabwe 1; and according to specialty: 17 all-rounders were chosen, 10 wicket-keepers, 25 bowlers and 23 batsmen. Prior to the first auction, the BCCI laid down a number of guidelines for each franchise which regulated the following: salary cap, the size and composition of each squad and the use of icon players. Since the teams had to effectively construct a squad of players, the auction can be classified as a multi-unit auction. Nevertheless, each franchise needed sufficient depth of playing talent in each of the game's specialisations: batting, bowling and wicket keeping.

Each franchise had a maximum of USD5 million to spend in the first players' auction, hence the auction also involved a budget constraint (there was also a minimum salary

cap of USD3 million, which proved to be non-binding). The amount of the winning bid (determined in an English-style setting) became the player's salary and this figure was included in the salary cap. Under-22 players from the BCCI were remunerated with a minimum annual salary of USD20,000, while the remaining players (Indian or foreign) were offered a base of USD50,000, effectively creating 'reserves' for these players. While each squad was restricted to a quota of eight foreign players, four at most could be in any given starting line up. The BCCI mandated that a minimum of four local players from both 'catchment areas' and the BCCI under-22 pool be included in each franchise. Despite the salary cap restriction of USD5 million, five franchises were allowed ultimately to exceed this figure (Chennai, Delhi, Hyderabad, Mohali and Mumbai), while Kolkata nudged USD6 million. Players unavailable for part of the tournament due to being on national duty (mainly Australian players) were paid on a pro-rata basis, thereby reducing the expenditure of franchises below USD5 million. Therefore, while the budget constraint proved to be binding (as it is in most professional sports where it is used), it could be argued that the league faced credibility problems in its enforcement.

The IPL conducted a secondary auction on 11 March, involving an additional list of 28 players who were available.[3] This included 14 from India's under-19 player list.[4] As noted above, the salary cap of USD5 million was a major constraint to at least half of the franchises. Jaipur was the only franchise to spend below the minimum threshold of USD3 million in the first auction, and hence were able to take full opportunity to add to its squad. IPL officials decided to relax the rule on overseas players and allow each franchise the option of picking up a ninth overseas player

[3] Only one player (James Hopes) sold for more than USD150,000 in the second auction. In the first auction, only three of the 80 players sold for less than this figure.
[4] These players were allocated according to a draft system.

because some players were unavailable for the entire season and two of the players contracted by the IPL before the first auction were still available.[5]

In summary, the various rules and restrictions, as well as the general dynamics and characteristics of the auction created a very unusual auction framework (much more so than the following open auction in January 2011), necessitating the consideration of numerous modelling issues. These issues are outlined in the following section.

4. Modelling Issues

4.1 Data Set

Our data set comprises the 80 players for whom the bidding price resulted in a 'sale'. The player characteristic and career statistic data were collated from the CricInfo website at: http://www.cricinfo.com/, the latter being compiled in March 2008 and backdated to 19 February wherever necessary. Since the volume of international Twenty20 matches played prior to February 2008 is low, player statistics from those matches would have been considered unreliable, forcing bidders to place more weight on career statistics from other forms of the game. Consequently, we estimate separate models using these statistics from Tests and ODIs (Karnik, 2010, ignores the former), as we believe bidders were well informed about players' abilities in both forms of the game in trying to assess their suitability for Twenty20 cricket. This makes sense, for Test cricket is a longer form of the game, and is hence likely to reveal more of a player's true qualities, whereas ODIs more closely resemble Twenty20 matches mechanically, in that both are shortened versions of the game with specific bowling and fielding restrictions in place and an emphasis on chasing run targets.

[5] Bangalore, Jaipur and Kolkata each exercised this option.

Selection of a narrow set of control variables and subsequent estimation via EBA is one possible alternative, as taken by Karnik (2010). However, we argue that such an approach creates the possibility of omitted variable bias, on the basis that the various bidders made use of a much broader set of information about the players involved. Admittedly, while many things are largely ignored, bidders do nonetheless have access to (and consider) many metrics regarding specific players when determining their valuations. To this end, we consider 57 potentially important explanatory variables. Their sample means, standard deviations, largest and smallest values are shown in columns 2-5 of table 3. In table 4, we assign the listed explanatory variables to a number of distinct groups, categorised by the broad athletic trait described by the statistic. The first two categories are based on identifiable characteristics (henceforth IC), and the remaining categories on career statistics (CS). This partition is displayed in column 6 of table 3. Note that it is not a one-to-one relation, as certain variables fall into more than one category. The variables highlighted below are also included in a separate appendix (see appendix). Since the aim of this research is to model player valuation in the IPL, the key variable is *VALUE*, the player's auction value in USD. We considered the following potentially important explanatory variables:[6]

PAGE: player's age in years on 20 February 2008;

NIND, *NAUS*, *NRSA*, *NSRL*, *NOTH*: dummy variables for the players' nationality (1 for Indian, Australian, South African, Sri Lankan and all other nationalities, respectively, and 0 otherwise);

SBAT, *SBWL*, *SWKT*, *SALR*: dummy variables for players regarded primarily as batsmen, bowlers, wicket-keepers and all-rounders, respectively;

[6] Discretionary judgement was required in constructing some binary variables, such as player specialisation and international team status variables, and for some players, the values of these variables may differ slightly between Tests and ODIs.

FLDR: a dummy variable for (outfield) players that have stand-out fielding ability;

XFTR: an 'X-Factor' dummy variable for players that have qualities that may generate extra value independent of playing ability (aura, looks, marketability, etc);

RETD: a dummy variable indicating the player had retired from international duties prior to the commencement of the IPL season;

IREG, *IYTD*, *IMSC*: dummy variables indicating the players are considered regular, yet-to-debut and miscellaneous (all other), respectively, in their respective international sides;

ICON: a dummy variable indicting that the player belongs to an IPL side that has an 'icon' player;

CAPT: a dummy variable indicating captaincy experience (defined as national side captaincy on at least two occasions);

FORM: a dummy variable indicating that a batsman/wicket-keeper (bowler) had a higher (lower) batting (bowling) average in 2007 than in his entire playing career – for all-rounders, both conditions are required;

MTCH: number of matches played;

BTIN, *RUNS*: the number of innings batted and runs scored;

HSCR: highest score;

BTAV: batting average;

BLSF: number of balls faced;

BTSR: batting strike rate;

C100, *H050*: the number of centuries and half-centuries scored;

FOUR, *SIXS*: the number of fours and sixes scored;

BWIN, *BLBD*, *RUNC*, *WKTS*: the number of innings bowled in, balls bowled, runs conceded and wickets taken, respectively;

BWAV, *BWSR*: the bowling average and strike rate, respectively; *ECON* is the economy rate;

05WI, *10WM*: the number of occasions a bowler took at least five or ten wickets in a (long-form) innings or in a match, respectively;

B4WK, *B5WK*: the number of occasions a bowler took at least four and five wickets in a (one-day) match;

TW20, *T20I*: the number of Twenty20 matches played at domestic or international levels, and at international level only.

Some of the playing statistics and variables are not uniform in different forms of the game. In these cases, superscripts indicate the relevant game form.

As regards the functional form of our hedonic regression model, we assume that it is log-lin, that is the dependent variable is the logarithm of *VALUE* but otherwise the model is linear in the independent variables, except *PAGE*. In this case, we allow for a quadratic relationship assuming that up to a certain point higher age implies player improvement/development and hence value, but thereafter it becomes a liability as inevitable athletic decline sets in.

Finally, coming back to table 3, the last column exhibits the expected signs of the relationships between *VALUE* and individual explanatory variables. While we have no expectation of most nationality and specialisation variables (nor *ICON* for that matter), we do expect to observe a premium for Indian players and perhaps all-rounders because of their versatility. We would expect to see premiums for international regulars, strong fielders and X-Factor players (see table 5 for a list of players assigned a value of one for the *FLDR* and *XFTR* dummies), form players,

players with captaincy experience, and those who have played more Twenty20 matches; while expecting discounts for retired players, or those yet-to-debut internationally. In terms of career statistics, all batting variables are defined such that 'more is better', therefore, we would expect to see positive estimates across the board; whereas with bowling, we would expect negative estimates only for statistics framed in terms of runs conceded per unit (specifically *RUNC*, *BWAV*, *ECON* and *BWSR*).

4.2 Model Specification and Estimation

Given the large number of potentially important explanatory variables, serious multicollinearity is most likely in a regression model containing all or most of them. In order to study this possibility, we calculated pairwise correlation coefficients between the possible pairs of independent variables and also between ln(*VALUE*) (henceforth *value*) and each independent variable.[7] We found 42 pairs of independent variables exhibiting extremely strong correlation, i.e. above 0.9 in absolute value, which is not surprising since some variables are by definition linear combinations of other variables. What is even more important, the highest correlation coefficient between *value* and any of the independent variables is only 0.345 (with *XFTR*) and close to one-quarter of all pairs of independent variables have stronger relationships.

In the light of these findings it was not surprising that our first regression with all possible independent variables could not be estimated due to near singularity. To avoid multicollinearity, we had to reduce the number of independent variables. Consequently, it was necessary to perform automatic variable selections with stepwise regressions. For each group of players (i.e. overall, batters, bowlers) and forms of the

[7] Given the enormity of the correlation matrix, it is not reported here but is available on request.

game (i.e Test and ODI) we experimented with the stepwise-forwards and swapwise selection methods. The stepwise-forwards procedure starts with a regression including just a constant term and then in each step augments the model with the variable that has the lowest *p*-value in the latest regression and removes each variable whose *p*-value is high, both compared to some stopping criterion. The swapwise selection method is similar to the stepwise-forwards method, but this time the decision rule to add a variable or to swap an 'inside' variable with an 'outside' variable (that is, a variable currently in the regression with one that is not) is based on the potential increment in R^2.[8] In four out of six cases the stepwise-forwards procedure led to reasonable and similar specifications to the swapwise method, but in the other two cases the swapwise results turned out to be superior intuitively and/or statistically.[9]

Our preferred models were subjected to four standard tests: (i) the *F*-test of overall significance, (F); (ii) the Jarque and Bera (1980) test for normality of the residuals, (N); (iii) White's (1980) heteroscedasticity (χ^2) test without cross terms, (W); and (iv) Ramsey's (1969) regression specification error test with two fitted terms $(RESET)$. Each model passed the *F*, *N* and *RESET* tests, but two failed the *W* test. In these latter cases we rely on White's heteroscedasticity-consistent standard errors. We also calculated the Durbin-Watson *d*-statistic (DW) for each preferred model from the residuals ordered according to increasing values of the dependent variable.

[8] Goodness-of-fit is reported in table 6 via both the standard (R^2) and adjusted $(\overline{R}^2)$.

[9] In the case of 'Overall ODI', the two regressions are the same, except that the stepwise-forwards procedure retained a statistically insignificant (at the 8% level) regressor. On the other hand, in the case of 'Bowling Test' the stepwise-forwards regression is far too parsimonious – it has only three regressors while the swapwise regression has 12 strongly significant and intuitively appealing regressors. Furthermore, its explanatory power is only about half of the swapwise regression (the adjusted-R^2 statistics are 0.33 and 0.63, respectively).

5. Results

Initially, two models were estimated for all 80 players purchased in the auction – one using Test (and Domestic) career statistics and the other using ODI statistics, as it is not clear *a priori* which set of statistics contain more information about the player's ability to perform in the IPL and hence which set bidders will give greater weight to. Since the volume of Twenty20 matches played up to February 2008 is thin, it is inadvisable to use career statistics in that form of the game, rather we use simply the number of games played, as any (albeit limited) experience at Twenty20 would be considered potentially valuable.

The results are summarised in table 6. Overall, the various models produce some intriguing results. With only one exception, all models explain over 60 per cent of the variation in *value*, well in excess of the fit obtained by Karnik (2010), and variables from most categories appear as significant in most models, with the exception of the FP category. The most frequently appearing variables are *NIND*, *XFTR* and *TW20*, each appearing in five models out of six, suggesting strong evidence of value premiums for Indian players, those with an X-Factor characteristic, and those with more experience in this specific form of the game. The home player bias is arguably the most predictable result (presuming that management decisions reflect perceived fan preferences), and is consistent with numerous other studies, such as Kanazawa and Funk (2001) and Foley and Smith (2007), but in direct contrast to the findings of Wilson and Ying (2003). Furthermore, virtually all significant (semi-elasticity) parameter estimates are of the anticipated sign, except for AC-category domestic variables, which may be attributable to non-Test regulars (generally less valuable

players) playing more domestic matches, allowing them to accumulate more 'contribution' at that level.

The Test Model results (ignoring factors mentioned already as being standard) indicate, *inter alia*, that players who had retired already from playing Tests were valued less, reflecting their reduced match conditioning. Players signed with teams containing an icon player also received higher bids. A player's propensity to hit sixes (less common in Tests) when batting is a strong positive indicator of the value of big-hitting ability. For bowling, ability to take wickets is valued at Test level, but the analogous domestic coefficient estimate is significantly negative, possibly due to interaction with the Test wickets term. The sign on bowling average is negative but positive for strike rate, suggesting that better bowlers are valued more highly, but combined with a higher strike rate means that such bowlers are also more economical.

For the ODI model, *PAGE* becomes significantly negative, consistent with the observations of commentators that Twenty20 is '…a young man's game', however, $PAGE^2$ is not significant. In addition to the Indian premium, there is also a premium for Australian and South African players reflecting recent strong performance of these international teams. Also, one-day experience in terms of number of matches played and number of sixes hit are also both significant.

With a mix of batting and bowling statistics evident in the full sample, the next question to address is that of whether more precise estimates can be obtained by splitting the sample into batting and bowling specialisations. For the former, we include only records for which either *SBAT*=1 or *SWKT*=1, on the assumption that

wicket-keepers are still heavily valued for their batting skills. Here, bowling variables are ignored completely, since most batsmen have bowled a small amount during their career, thus possibly creating distortions in the results if they were left in. For this reason, all-rounders are excluded since they are valued heavily on their bowling as well. However, for bowling regressions (requiring either *SBWL*=1 or *SALR*=1 for inclusion), batting variables are still included. The reason for this is that even a specialist bowler may still have some additional value if they have an ability to play a 'cameo' role when they are required to bat with a few overs still remaining in an innings.

For the selected batting group, the Test model is really the only model in which most of the significant variables are career statistics as opposed to identifiable characteristics. Total test runs scored proves to be significant, but so are both domestic average and number of centuries scored, covering elements of categories *AB*, *AC* and even *FP* for batsmen. Interestingly, the identifiable characteristics take over in the ODI model for batsmen. Fielding prowess becomes significant (thought to make a larger difference to the outcome in shorter forms of the game), as well as experience in the form of number of ODI appearances.

In the models for bowlers, some batting variables prove to still be significant. For the Test model, an Australian premium arises once again (perhaps due largely to Shane Warne and Glenn McGrath), while players retired from or yet to play tests are valued at a discount. Like before, two significant domestic variables (matches and strike rate) are not the anticipated sign, whereas $BWSR^{T}$ and $05WI^{T}$ demonstrate the value of taking wickets more frequently and the ability to take big hauls. When batting,

bowlers who have greater ability are valued higher, even though their ability to score quickly is clearly not, as evidenced by the signs on the $BTAV^T$ and $BTSR^T$ terms.

With respect to diagnostics, it is seen in table 6 that there is some evidence of positive serial correlation (based on increasing *value*). While this typically implies model misspecification, we argue that there is an underlying interpretation – that of bidder irrationality. Figure 1 plots a solid line of (increasing) fitted values according to the geometric mean of USD player values from each of the models (this will be an average of four of the six models for each player).[10] We identify and label three outliers outside the approximate 95 per cent confidence interval – Mohammad Kaif and Kumar Sangakkara above, and Scott Styris below.[11]

While not so evident in the averages, there nevertheless appears to be some evidence of underbidding (overbidding) at the bottom (top) end of the player pool. This could be demonstrative of a Winner's Curse for 'star' players, especially if they *ex-post* underperform in the tournament – this is supported anecdotally by Rajasthan, who won the inaugural tournament, despite spending the least of all teams in the first round auction. When players are ordered instead with respect to actual valuations, we find that 28 out of the 40 less expensive players were underbid according to our mean valuations, whereas only 14 of the more expensive 40 players were underbid (this difference is significant at the one per cent level of a χ^2-distributed difference of means test). At a cursory glance, figure 1 also seems to suggest that the greatest absolute percentage valuation errors occur in the mid-range of the market.

[10] The use of a mean value for this comparative exercise is on the grounds that the various models differ quite markedly, and do not always include the same sample of players.
[11] Shahid Afridi is also labelled and worthy of a special mention – he is just inside the lower-bound interval, and a bigger outlier than the other three in terms of absolute (as opposed to logarithmic) value.

Such behaviour in the sports industry is not unheard of, due to the existence of 'superstar effects' – see Matthews, Sommers and Peschiera (2007) for an illustration in women's golf with respect to the distribution of prize money in tournaments. While in this particular case, our sample is also drawn from the upper-tail of the population, the outcome may be driven largely by the various restrictions placed on the auction referred to earlier, as the true valuations of players cannot be revealed completely within this framework. Specifically, it is likely that the budget constraint may have limited expenditures on the 'lesser-lights', since bidding for 'star' players, for which bidding was most highly competitive, was concentrated heavily in the earlier rounds of the auction.

An associated issue is that of sequencing of players within the auction, and whether players' values were affected by the round in which they were auctioned. Unfortunately, due to the nature of how the players were allocated to rounds and the way the rounds were sequenced (see section 3), there is considerable correlation with playing specialisations – a factor that we believe to be more crucial. Furthermore, many of the players expected to command higher prices were allocated to earlier rounds. Therefore, it is not suitable to introduce the auction round as a control variable specifically. However, it is still possible to examine this effect indirectly. Table 7 shows how the mean residual from the regressions (interpreted in this context as player over- or under-valuation) evolves over the rounds of bidding. The rounds are aggregated as 1-3, 4-6 and 7-8 to reduce noise in the means. We would expect the icon players (whose prices were not determined directly) to be overvalued on average because of their premium status. Further down, we see that the mean value does not

decline (nor rise) monotonically throughout the rounds of the auction (circumventing concern that auction round was omitted incorrectly). Nevertheless, there is still some interest in these results: there was a tendency for players to be overvalued in the first few rounds, to be underbid in the middle rounds, and to be overvalued in the latter rounds. Players that were relegated to the reserve pool were undervalued on average as would be expected, but only very slightly.

Finally, in contrasting the use of Test versus ODI models, when the full sample is used, it is difficult to separate the models purely on the basis of the measures of fit or information criteria. However, in the player specialisation decompositions, we see an interesting development. The fit and information criteria measures show that for batsmen, ODI stats may be given slightly more weighting in assessment, whereas Test information is given more weighting for bowlers. Such an observation is consistent with the apparent consensus of cricket commentators, who note that the scope for a single bowler to determine the outcome of a Twenty20 match is limited compared to that of a single batsman. Thus, perhaps ODI data is more useful in assessing batters because of the similarity of the formats, whereas assessors fall back on Test data for bowlers as a general indication of true 'quality'.

6. Conclusion

We have undertaken a formal evaluation of the determinants of the marginal revenue product of professional cricketers, as established by an auction with a very unique range of characteristics that can allow us to make inferences on the effect of such restrictions on price. The problem is a very appealing one for numerous reasons, among them: (i) the IPL is arguably the biggest revolutionary concept in the game's

history; (ii) the auction is a very rare opportunity to measure athletes' valuations in this way; (iii) the free availability of player data.

Our most compelling findings are as follows: (i) our six models have strong explanatory power for cross-sectional models; (ii) variables from most defined categories appear as significant in most models; (iii) most of our significant estimates are of the anticipated sign. The variables most commonly appearing to have a material bearing on player value are the existence of Indian-player and X-Factor premiums, as well as a positive relation with previous Twenty20 experience (number of games). Furthermore, we interpret the presence of serial correlation as some evidence of overbidding for star players and underbidding for lesser players. Finally, ODI statistics seem to provide more informational content about batsmen than Test statistics, while the inverse is true for bowlers.

There is great potential for future work in auction theory to model this auction (with its unique set of attributes) in a formal setting, with a view to considering whether the findings are consistent with those presented here, as well as those from classic works in sports economics, most notably the seminal Fort and Quirk (1995) model. Nevertheless, while not comparable directly with results from any future auctions, the results could be used by the various IPL franchises, along with other information, to approximate 'fair' valuations for players in future seasons, to mitigate the incidence of overbidding.

Table 1: Number of International ICC-Sanctioned Cricket Matches by Type, 1993-2009 (Not Counting Cancelled or Abandoned Matches)

Calendar Year	Tests	ODIs	Twenty20
1993	36	82	0
1994	38	98	0
1995	40	60	0
1996	28	127	0
1997	44	115	0
1998	45	108	0
1999	43	154	0
2000	46	131	0
2001	55	120	0
2002	54	145	0
2003	44	147	0
2004	51	128	0
2005	49	107	3
2006	46	160	9
2007	31	191	38
2008	47	126	29
2009	41	123	49
2010	43	152	60

Source: http://www.cricinfo.com/

Table 2: Number of Domestic-Level Matches by Type in All 10 Test-Playing Countries, 2000-2008 (Not Counting Cancelled or Abandoned Matches)

Calendar Year	Long-Form	Limited Overs	Twenty20
2001	773	665	0
2002	693	847	0
2003	731	669	48
2004	633	649	70
2005	604	626	159
2006	598	584	168
2007	680	615	236
2008	615	612	262

Source: http://www.cricinfo.com/

Table 3: Descriptive Statistics

Variable	Mean	Std. Dev.	Maximum	Minimum	Category	Sign
ln(*value*)	12.983	0.647	14.221	11.513	NA	
PAGE	28.700	4.811	38.000	19.000	PC	*
$PAGE^2$	846.55	278.15	1,444.0	361.00	PC	-
NIND	0.363	0.484	1.000	0.000	PC	+
NAUS	0.163	0.371	1.000	0.000	PC	?
NRSA	0.125	0.333	1.000	0.000	PC	?
NSRL	0.138	0.347	1.000	0.000	PC	?
NOTH	0.213	0.412	1.000	0.000	PC	?
$SBAT^T$	0.388	0.490	1.000	0.000	CC	?
$SBAT^O$	0.350	0.480	1.000	0.000		
$SBWL^T$	0.338	0.476	1.000	0.000	CC	?
$SBWL^O$	0.338	0.476	1.000	0.000		
SWKT	0.113	0.318	1.000	0.000	CC	?
$SALR^T$	0.163	0.371	1.000	0.000	CC	+
$SALR^O$	0.200	0.403	1.000	0.000		
$FLDR^T$	0.100	0.302	1.000	0.000	CC	+
$FLDR^O$	0.113	0.318	1.000	0.000		
$XFTR^T$	0.113	0.318	1.000	0.000	CC,PC	+
$XFTR^O$	0.138	0.347	1.000	0.000		
$RETD^T$	0.088	0.284	1.000	0.000	CC	-
$RETD^O$	0.088	0.284	1.000	0.000		
$IREG^T$	0.475	0.503	1.000	0.000	CC	+
$IREG^O$	0.550	0.501	1.000	0.000		
$IYTD^T$	0.125	0.333	1.000	0.000	CC	-
$IYTD^O$	0.025	0.157	1.000	0.000		
$IMSC^T$	0.313	0.466	1.000	0.000	CC	0
$IMSC^O$	0.338	0.476	1.000	0.000		
ICON	0.625	0.487	1.000	0.000	NA	?
$CAPT^T$	0.225	0.420	1.000	0.000	CC	+
$CAPT^O$	0.250	0.436	1.000	0.000		
$FORM^T$	0.313	0.466	1.000	0.000	CC	+
$FORM^O$	0.350	0.480	1.000	0.000		
$MTCH^T$	47.913	43.098	146.00	0.000	EX	+
$MTCH^O$	132.23	107.48	413.00	0.000		
$BTIN^T$	73.938	67.085	237.00	0.000	EX	+
$BTIN^D$	191.05	121.08	571.00	18.000		
$RUNS^T$	2,416.9	2,937.5	11,782	0.000	AC	+
$RUNS^D$	6,548.9	6,149.1	26,277	58.000		
$RUNS^O$	2,752.7	3,531.6	16,088	0.000		
$HSCR^T$	127.95	103.68	380.00	0.000	FP	+
$HSCR^D$	176.81	95.539	380.00	23.000		
$HSCR^O$	86.350	57.602	189.00	0.000		
$BTAV^T$	26.653	18.668	78.143	0.000	AB	+
$BTAV^D$	33.823	15.145	60.059	7.273		
$BTAV^O$	23.735	13.885	55.429	0.000		

BLSF^{T}	4,207.0	5,444.0	23,582	0.000	AB,EX	+
BLSF^{O}	3,437.4	4,397.4	18,839	0.000		
BTSR^{T}	43.479	21.695	86.131	0.000	EE	+
BTSR^{O}	72.325	21.726	116.67	0.000		
C100^{T}	5.488	8.853	39.000	0.000	AB,AC,FP	+
C100^{D}	15.625	19.543	80.000	0.000		
C100^{O}	3.888	7.314	41.000	0.000		
H050^{T}	10.963	14.009	51.000	0.000	AB,AC,FP	+
H050^{D}	30.200	28.666	99.000	0.000		
H050^{O}	15.325	21.735	87.000	0.000		
FOUR^{T}	284.91	341.49	1,245.0	0.000	AB,AC	+
FOUR^{O}	265.51	369.11	1,758.0	0.000		
SIXS^{T}	15.600	20.541	100.00	0.000	AB,AC	+
SIXS^{O}	37.575	54.026	247.00	0.000		
BWIN	45.213	65.611	273.00	0.000	EX	+
BLBD^{T}	4,795.8	9,167.6	40,705	0.000	AB,EX	+
BLBD^{D}	10,787	14,874	74,830	0.000		
BLBD^{O}	3,341.9	4,520.7	16,364	0.000		
RUNC^{T}	2,261.7	4,028.5	17,995	0.000	AC	-
RUNC^{D}	5,183.2	6,611.4	34,449	0.000		
RUNC^{O}	2,522.7	3,239.9	11,206	0.000		
WKTS^{T}	78.713	158.65	723.00	0.000	AC	+
WKTS^{D}	191.11	275.42	1,319.0	0.000		
WKTS^{O}	85.663	118.01	459.00	0.000		
BWAV^{T}	27.899	24.480	105.63	0.000	AB	-
BWAV^{D}	34.132	23.918	187.00	0.000		
BWAV^{O}	26.709	25.429	151.00	0.000		
ECON^{T}	2.602	1.837	13.000	0.000	EE	-
ECON^{D}	3.146	1.024	7.588	0.000		
ECON^{O}	4.078	2.988	18.000	0.000		
BWSR^{T}	54.195	45.565	210.00	0.000	AB	-
BWSR^{D}	65.304	42.851	330.00	0.000		
BWSR^{O}	32.186	28.541	150.00	0.000		
05WI^{T}	3.825	9.900	62.000	0.000	AB,AC,FP	+
05WI^{D}	9.063	18.047	114.00	0.000		
10WM^{T}	0.775	2.719	20.000	0.000	AB,AC,FP	+
10WM^{D}	1.613	4.496	32.000	0.000		
B4WK	2.125	3.421	13.000	0.000	AB,AC,FP	+
B5WK	0.938	1.716	8.000	0.000	AB,AC,FP	+
TW20	13.975	9.445	45.000	0.000	FY	+
T20I	5.838	4.465	15.000	0.000	FY	+

Note: the sample size is 80 for all variables. For playing statistics and variables not uniform in different forms of the game, ^{T} indicates Tests; ^{D} indicates first-class/domestic (long-form); and ^{O} indicates ODIs. *Sign* refers to the expected sign of the relationship between *value* and a particular independent variable, (*) indicates expected sign to be '+' if $PAGE^2$ enters the model, and '-' if it does not.

Table 4: Categories to Which Variables are Assigned

Identifiable Characteristics (IC)	Code	Career Statistics (CS)	Code
Personal characteristics	PC	Ability (mean-driven)	AB
Cricketing characteristics	CC	Experience	EX
		Accumulated contribution	AC
		Expeditiousness/economy	EE
		Freak performances	FP
		Familiarity	FY

Note: 'freak performances' is the possibility that such performances create a 'halo effect' in the minds of bidders.

Table 5: Players Assigned Value of One for *XFTR* and *FLDR* Dummies

Fielding		
Player	National Team	IPL Team
AB de Villiers	South Africa	Delhi
Dilshan Tillakaratne	Sri Lanka	Delhi
Andrew Symonds	Australia	Hyderabad
Shahid Afridi	Pakistan	Hyderabad
Herschelle Gibbs	South Africa	Hyderabad
Mohammad Kaif	Pakistan	Jaipur
Ricky Ponting	Australia	Kolkata
Yuvraj Singh	India	Mohali
Suresh Raina	India	Chennai
X-Factor		
Player	National Team	IPL Team
Mahendra Dhoni	India	Chennai
Andrew Symonds	Australia	Hyderabad
Adam Gilchrist	Australia	Hyderabad
Shahid Afridi	Pakistan	Hyderabad
Shane Warne	Australia	Jaipur
Ishant Sharma	India	Kolkata
Shoaib Akhtar	Pakistan	Kolkata
Brett Lee	Australia	Mohali
Sanath Jayasuriya	Sri Lanka	Mumbai

Note: Suresh Raina was yet to make his test debut at the time of the IPL auction and is only given this status at ODI level

Table 6: Regression Results for Tests and ODIs

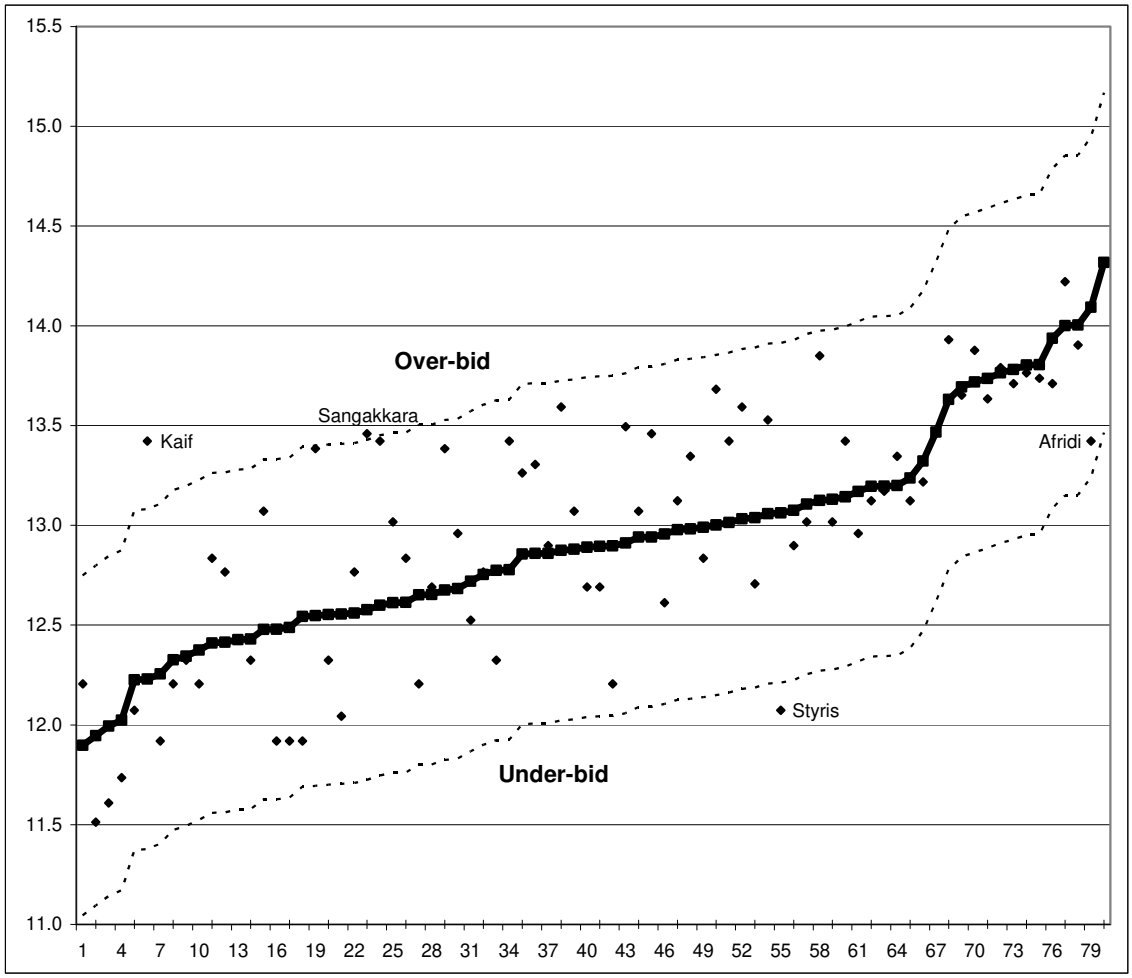
	Overall Test	Overall ODI	Batting Test	Batting ODI	Bowling Test	Bowling ODI
C	11.660[a]	13.421[a]	10.838[a]	11.705[a]	12.229[a]	12.816[a]
PAGE		-0.064[a]				
NIND	0.706[a]	0.854[a]	0.816[a]	1.006[a]	0.577[a]	
NAUS		0.552[a]			0.557[a]	
NRSA		0.349[c]				
NSRL						-0.716[a]
NOTH						-0.415[b]
FLDR				0.438[b]		
XFTR	0.848[a]	0.698[a]		0.731[b]	0.600[a]	0.406[b]
RETD	-0.729[a]				-0.570[c]	
IREG						0.388[b]
IYTD					-1.671[a]	
ICON	0.213[b]					
$MTCH^O$		0.004[a]		0.003[a]		
$MTCH^D$					-0.005[a]	
$RUNS^T$			9.1×10^{-5}[b]			
$BTAV^T$					0.035[a]	
$BTAV^D$			0.036[a]			
$BTSR^T$					-0.018[a]	
$C100^D$			-0.030[a]			
$FOUR^O$						0.001[b]
$SIXS^T$	0.012[a]					
$SIXS^O$		0.015[a]				
$WKTS^T$	0.004[a]					
$WKTS^D$	-0.002[a]					
$BWAV^D$	-0.032[a]					
$BWSR^T$					-0.016[a]	
$BWSR^D$	0.021[a]				0.030[a]	
$05WI^T$					0.036[a]	
TW20	0.036[a]	0.021[a]	0.025[a]	0.029[a]	0.026[b]	
R^2	0.630	0.607	0.626	0.661	0.747	0.528
$\overline{R}^2$	0.577	0.569	0.558	0.606	0.634	0.464
AIC	1.235	1.221	1.435	1.311	1.095	1.374
SC	1.562	1.460	1.731	1.573	1.644	1.620
F	11.770[a]	15.876[a]	9.193[a]	12.076[a]	6.639[a]	8.276[a]
N	0.570	0.955	0.629	0.556	0.726	0.226
W	9.632	16.152[a]	3.658	2.946	23.778[b]	6.536
RESET	0.304	0.286	2.072	0.147	1.465	1.134
DW	1.194[a]	1.307[b]	1.389[b]	2.539	0.938[a]	0.999[a]
n	80	80	40	37	40	43
Method	*SF*	*SWMIN*	*SF*	*SF*	*SWMAX*	*SF*

Note: [a], [b] and [c] indicate significance at 1%, 5% and 10% levels based on two-tail *t*-tests. *AIC* and *SC* denote the Akaike (1973 and 1977) and Schwarz (1978) Information Criterion, respectively; and *n* is the sample size. *SF* refers to stepwise-forwards regression with $p = 0.1$ stopping criterion for both forwards and backwards. *SWMIN* and *SWMAX* refer to swapwise regressions based on min and max R^2 increment, respectively.

Table 7: Mean Residual from Regressions According to Round of Bidding

Round(s)	Mean Residual	Sub-Sample
Icon	0.1963	5
1-3	0.1671	17
4-6	-0.0485	26
7-8	0.0902	24
Reserve	-0.0276	8

Figure 1: Rank-Fitted (Thick Line) with Approximate 95 Per Cent Confidence Intervals (Dashed Line) and Actual (Scatter Plots) Auction Values



References

Akaike, H. (1973), "Information Theory and an Extension of the Maximum Likelihood Principle", in *2nd International Symposium on Information Theory*, eds Petrov, B. N. and Csaki, F., Akadémiai Kiadó, Budapest, pp. 267-281.

Akaike, H. (1977), "On Entropy Maximization Principle", in *Applications of Statistics*, ed. Krishniah, P. R., North Holland, Amsterdam, pp. 27-41.

Allsop, P. E. and Clarke, S. R. (2004) "Rating Teams and Analysing Outcomes in One-Day and Test Cricket", *Journal of the Royal Statistical Society (Series A)*, 167 (4), 657-667.

Berri, D. J., Schmidt, M. B. and Brook, S. L. (2006), *The Wages of Wins: Taking Measure of the Many Myths in Modern Sport*, Stanford University Press, Stanford.

Bhattacharya, M. and Smyth, R. (2003), "The Game is Not the Same: The Demand for Test Match Cricket in Australia", *Australian Economic Papers*, 42 (1), 77-90.

Bodvarsson, O. B. and Brastow, R. T. (1998), "Do Employers Pay for Consistent Performance?: Evidence from the NBA", *Economic Inquiry*, 36 (1), 145-160.

Brooks, R. D., Faff, R. W. and Sokulsky, D. (2002) "An Ordered Response Model of Test Cricket Performance", *Applied Economics*, 34 (18), 2353-2365.

Downward, P., Dawson, A., and Dejonghe, T. (2009), *Sports Economics: Theory, Evidence and Policy*, Elsevier/Butterworth-Heinemann, London.

Ehenberg, R. G. and Bognanno, M. L. (1990), "Do Tournaments Have Incentive Effects?", *Journal of Political Economy*, 98 (6), 1307-1324.

Foley, M. and Smith, F. H. (2007), "Consumer Discrimination in Professional Sports: New Evidence from Major League Baseball", *Applied Economics Letters*, 14 (13), 951-955.

Fort, R. and Quirk, J. (1995), "Cross-Subsidisation, Incentives, and Outcomes in Professional Team Sports Leagues", *Journal of Economic Literature*, 33 (3), 1265-1299.

Frick, B. (2003), "Contest Theory and Sport", *Oxford Review of Economic Policy*, 19 (4), 512-529.

Frick, B. (2007), "The Football Players' Labor Market: Empirical Evidence from the Major European Leagues", *Scottish Journal of Political Economy*, 54 (3), 422-446.

Goddard, J. and Wilson, J. O. S. (2009), "Racial Discrimination in English Professional Football: Evidence from an Empirical Analysis of Players' Career Progression", *Cambridge Journal of Economics*, 33 (2), 295-316.

Hill, J. R. (2004), "Pay Discrimination in the NBA Revisited", *Quarterly Journal of Business and Economics*, 43 (1-2), 81-92.

Jarque, C. M. and Bera, A. K. (1980), "Efficient Test for Normality, Heteroscedasticity and Serial Independence of Regression Residuals", *Economics Letters*, 6 (3), 255-259.

Kanazawa, M. T. and Funk, J. P. (2001), "Racial Discrimination in Professional Basketball: Evidence from Nielsen Ratings", *Economic Inquiry*, 39 (4), 599-608.

Karnik, A. (2010), "Valuing Cricketers Using Hedonic Price Models", *Journal of Sports Economics*, 11 (4), 456-469.

Khan, L. M. (2000), "The Sports Business as a Labour Market Laboratory", *Journal of Economic Perspectives*, 14 (3), 75-94.

Leeds, M. A. (1988), "Rank-Order Tournaments and Worker Incentives", *Atlantic Economic Journal*, 16 (2), 74-77.

Lenten, L. J. A. (2008), "Is the Decline in the Frequency of Draws in Test Match Cricket Detrimental to the Long Form of the Game?", *Economic Papers*, 27 (4), 364-380.

Lewis, M. M. (2003), *Moneyball: The Art of Winning an Unfair Game*, W.W. Norton, New York.

Matthews, P. H., Sommers, P. M. and Peschiera, F. J. (2007), "Incentives and Superstars on the LPGA Tour", *Applied Economics*, 39 (1-3), 87-94.

Parsons, C. and Smith, I. (2007), "The Price of Thoroughbred Yearlings in Britain", *Journal of Sports Economics*, 9 (1), 43-66.

Preston, I. (2006), "The Economics of Cricket", in *Handbook on the Economics of Sport,* eds Andreff, W. and Szymanski, S., Edward Elgar, Cheltenham and Northampton, pp. 585-593.

Ramsey, J. B. (1969), "Tests for Specification Errors in Classical Least-Squares Regression Analysis", *Journal of the Royal Statistical Society (Series B)*, 31 (2), 350-371.

Rosen, S. and Sanderson, A. (2001), "Labour Markets in Professional Sports", *Economic Journal*, 111 (469), F47-F68.

Schwarz, G. (1978), "Estimating the Dimension of a Model", *Annals of Statistics*, 6 (2), 461-464.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48 (4), 817-838.

Wilson, D. P. and Ying, Y.-H. (2003), "Nationality Preferences for Labour in the International Football Industry", *Applied Economics*, 35 (14), 1551-1559.

Appendix

Dummy Variables

Code	Description
NIND	Indian
NAUS	Australian
NRSA	South African
NSRL	Sri Lankan
NOTH	All other nationalities
SBAT	Player regarded primarily as a batsman
SBWL	Player regarded primarily as a bowler
SWKT	Player regarded primarily as a wicketkeeper
SALR	Player regarded primarily as an all-rounder
FLDR	Outfield player with stand-out fielding ability
XFTR	Player with qualities that may generate extra value independent of playing ability: aura, looks, marketability, etc
RETD	Player retired from international duties prior to the IPL season
IREG	Player is considered a regular in his international side
IMSC	Player is considered a miscellaneous player in his international side
IYTD	Player is yet to debut in his international side
ICON	Player belongs to an IPL side that has an 'icon' player
CAPT	Player has captaincy experience (captained national side at least twice)
FORM	Batsman/wicket-keeper (bowler) had a higher (lower) batting (bowling) average in 2007 than in his entire playing career – for all-rounders, both conditions are required

Batting Variables

Code	Description
BTIN	Number of innings batted
RUNS	Number of runs scored
HSCR	Highest score
BTAV	Batting average
BLSF	Number of balls faced
BTSR	Batting strike rate
C100	Number of centuries scored
H050	Number of half-centuries scored
FOUR	Number of fours scored
SIXS	Number of sixes scored

Bowling Variables

Code	Description
BWIN	Number of innings bowled in
BLBD	Number of balls bowled
RUNC	Number of runs conceded
WKTS	Number of wickets taken
BWAV	Bowling average
BWSR	Bowling strike rate
ECON	Economy rate
05WI	Number of times bowler took at least five wickets in a long-form innings
10WM	Number of times bowler took at least ten wickets in a (long-form) match
B4WK	Number of times bowler took at least four wickets in an ODI match
B5WK	Number of times bowler took at least five wickets in an ODI match

Level of Twenty20 Cricket

Code	Description
TW20	Number of Twenty20 matches played at domestic level
T20I	Number of Twenty20 matches played at international level