Sale price as a component of life cycle costing for second hand platforms

Depreciation modelling in the context of the FFG-7 frigate

Fred Ma
Materiel Group Operational Research

DRDC CORA TM 2013–227
December 2013

Defence R&D Canada
Centre for Operational Research and Analysis

Materiel Group Operational Research
Assistant Deputy Minister (Materiel)

Sale price as a component of life cycle costing for second hand platforms

Depreciation modelling in the context of the FFG-7 frigate

Fred Ma
Materiel Group Operational Research

Defence R&D Canada – CORA
Technical Memorandum
DRDC CORA TM 2013–227
December 2013

Principal Author

Fred Ma

Approved by

R. M. H. Burton
Section Head, Joint Systems Analysis

Approved for release by

P. Comeau
Chief Scientist, DRDC CORA

Abstract

In Q3 of 2012, as a subtask within a Director Materiel Policy and Procedures Major Equipment Procurement Study (New versus Used), Directorate Materiel Group Operational Research (DMGOR) planned a study into the life cycle costing of second hand platforms (vehicles, ships, aircraft). As an initial input into that study, DMGOR undertook the estimation of the depreciation of two classes of frigates, the U.S. FFG-7 and the Dutch Kortenaer, from second hand sales prices. That work is reported here. A database of both frigate classes was developed from open source, but new and used price data were readily available only for the FFG-7s. Depreciation fit an exponential decay model, with an average loss of 8.4%/year and a 68% (±1 sigma) confidence interval of [5.9%, 11.1%]/year. One variable, vessel age, explained up to 56% of the second used sale price data. The data was not sufficient to extract the portion of depreciation due to aging design/technology rather than physical aging of the platform, but neither did it contradict published cost growth trend of 2%/year for buying new vessels of this type.

Because of the projected growth in real cost for new defence platforms, a thorough analysis of procurement options is becoming increasingly important, including the life cycle costing of buying used. This is an essential part of comparing with buying new that ideally includes platform suitability, proficiency, and timeliness of full operational capability. Existing cost models within DMGOR can potentially be developed for buying used. Exploration of this approach entails examination of disruptive effects due to change of ownership, the availability of data, the effect of aging design/technology on the estimation of depreciation, and the effect of value enhancing accessories such as sensor/weapons systems, accompanying aircraft, and support services. Methods developed for analyzing second hand purchases can potentially be adapted for other major platforms, including other maritime vessels, vehicles, and aircraft. They can also help situate the price range for selling assets, if the Canadian Armed Forces choose to consider this.

Résumé

Au troisième trimestre de l'année 2012, le Directeur - Recherche opérationnelle (Groupe des matériels) (DROGM) a entrepris une étude sur l'établissement du coût du cycle de vie des plateformes de seconde main (véhicules, navires, aéronefs). Cette étude était une sous-tâche de l'étude sur l'acquisition d'équipement majeur neuf et usagé du Directeur - Politiques et procédures (Matériel). Pour commencer cette étude, le DROGM a fait une estimation de la dépréciation de deux classes de frégates - la Kortenaer des Pays Bas et la FFG-7 des États Unis - à partir de leurs prix de vente à l'état usagé. Cette estimation est présentée dans le rapport. On a monté une base de données pour les deux classes de frégates en cherchant dans des sources ouvertes. Cependant, il a été possible de trouver les prix à l'état neuf et à l'état usagé seulement pour la FFG-7. La dépréciation a été calculée à l'aide d'un modèle de décroissance exponentielle, et on a déterminé qu'elle s'établissait en moyenne à 8.4% par année, avec 68% (±1 sigma) d'intervalle de confiance de [5.9%, 11.1%] par année. L'une des variables, soit l'âge du navire, expliquait jusqu'à 56% du prix de vente à l'état usagé. Les données n'étaient pas suffisantes pour déterminer dans quelle mesure la dépréciation était due à l'âge du modèle et de la technologie par rapport au vieillissement physique de la plateforme, mais rien ne réfutait l'augmentation annuelle de 2% du coût d'achat de navires neufs du même type.

Vu l’augmentation prévue du coût réel des plateformes de défense neuves, il devient de plus en plus important d’effectuer une analyse en profondeur des différentes options d’achat, ce qui implique notamment de déterminer le coût du cycle de vie d’une plateforme usagée. Il s’agit là d’un élément essentiel lorsqu’on compare l’achat d’équipement neuf à l’achat d’équipement usagé, outre la pertinence, l’aptitude et les délais dans lesquels la capacité opérationnelle totale sera atteinte. Les modèles de coût utilisés par le DROGM pourraient être appliqués à l’achat d’équipement usagé. L’exploration de cette approche implique d’examiner les effets perturbateurs du changement de propriétaire, la disponibilité des données, les effets de l’âge du modèle et de la technologie sur l’estimation de la dépréciation, et les effets des éléments augmentant la valeur de la plateforme (p. ex., capteurs, systèmes d’armes, aéronef de soutien, et services de soutien). Les méthodes créées pour analyser les achats d’équipement de seconde main pourraient être adaptées pour l’analyse d’autres plateformes majeures (autres navires, véhicules et aéronefs). Elles peuvent aider à déterminer la fourchette de prix pour la vente d’équipement dans l’éventualité où les Forces armées canadiennes décidaient d’envisager cette possibilité.

Executive summary

Sale price as a component of life cycle costing for second hand platforms : Depreciation modelling in the context of the FFG-7 frigate

Fred Ma; DRDC CORA TM 2013–227; Defence R&D Canada – CORA; December 2013.

Background: In May 2012, Director Materiel Policy and Procedures Major Equipment Procurement initiated a study into the trade-offs of purchasing new versus used major equipment, of which cost analysis is a major component. Quantitative methods for analyzing costs associated with buying used can be applied to various platforms, such as frigates, support ships, destroyers, vehicles, and aircraft. Such a financial analysis is an essential part of a comprehensive comparison that ideally encompasses additional factors such as operational suitability of candidate second hand platforms, i.e., do they have the required capabilities, are they proficient enough in those capabilities, and the timeliness with which full operational capability (FOC) can be reached[1].

Objectives: This report describes the preliminary planning of life cycle cost modelling for buying second hand major platforms, and a case study in the forecasting of one major cost component, sale price, for an example platform, the FFG-7 frigate. In this report, we treat platform depreciation as synonymous with expected second hand sale price[2].

Method: Factors potentially affecting a platform's second hand life cycle cost pattern were identified in the preliminary planning. A subset of these factors was used to define a small database with which depreciation could be analyzed as a function of platform age. The main factors are age-related (new and second hand purchase dates) and costs (new and second hand). Two naval platforms were initially selected as candidates for which there were reasonable prospects of obtaining the data: the Dutch Kortenaer and the US FFG-7. An open source data gathering phase was undertaken to populate the database (new and used cost data was readily available only for the FFG-7), followed by application of a simple exponential decay model for constant-dollar depreciation.

Method Limitations: A major limitation in this study was the difficulty in obtaining new and used cost information for platforms that have been sold second hand. Furthermore, the second hand purchases in this study are well in the past. In contrast, the second hand platforms on which the Canadian Armed Forces (CAF) may need advice are future purchases. There may not be many past instances of second hand sales of such platforms, which would exacerbate the limited availability of data.

[1] Including anticipated restorations, Canadianizations (or more generally, modifications to meet requirements specific to the buying nation), and realistic time margins for unforeseen restorations/adaptions based on historical precedents. Cases that exemplify the importance of these considerations are cited in this report.

[2] We associate depreciation with used sales. The report discusses the distinction between depreciation versus the overall diminishment in value with time that can characterize a particular platform.

Modelling depreciation addresses a single component of life cycle cost, which itself is only one aspect of evaluating a purchase option, as indicated above. Moreover, the analysis reported here models one determinant of depreciation (age), albeit the main one. Finally, used sales represent only a portion of a platform's population. Knowledge of a platform's overall characteristic degradation with age would allow for a more complete assessment of buying used, including risks and required vigilance in terms of the warranted research and the depth of information to seek on the condition of the platform. The question of the best metric or proxy for overall degradation cannot be divorced from considerations of data availability.

Results: A database for the two classes of frigates was developed from open source information with the assistance of DHH, CDI's OSINT staff, and U.S. DSCA[3]. As mentioned, new and used cost data was only readily available for the FFG-7s. After accounting for inflation, the real depreciation was modelled as an exponential decay in second hand cost and found to be 8.4%/year, with a 68% confidence interval of [5.9%, 11.1%]/year. An attempt was made to estimate the portion of depreciation that can be attributed to the aging of platform design and technology relative to evolving military needs, but the sample was too small and spanned too narrow a time window. However, no evidence was found in our data to contradict a reported trend of 2% growth in real cost for applicable vessel classes bought new.

Theoretical developments in model validation are reported in the annexes.

Significance: The 2%/year growth in real cost for new purchases is actually at the low end of the range of growth rates reported for defence equipment costs. Annualized growths of 5 to 6% are not uncommon, and selected equipment classes can exceed 10%. With this new cost growth comes an increasing imperative to ensure that procurement options are analyzed as thoroughly as possible to maximize the confidence with which life cycle costs can be forecasted, including for options of buying second hand. A simple exponential decay model for depreciation was able to explain 41% to 56%[4] of the used sale price data using just one input variable (vessel age).

Future work: The input to any model is data, with higher fidelity modelling generally requiring more data. The forecasting of used purchase cost can leverage information beyond just platform age. Indeed, using many variables which potentially affect depreciation enables the estimation of used purchase cost for platforms other those for which data is gathered. More data is needed for such higher dimension modelling, but this approach has been demonstrated in the estimation of new purchase costs (the Joint Support Ship).

To improve estimation of second hand sale price, it would be useful to establish approaches to properly account for the value enhancements and supplementary equipment that accompany platforms that are sold used, which increase the variance in depreciation estimates based on used sale price. It may be possible to capture some of this dependence by augmenting a general multidimensional model with additional, judiciously defined input variables.

[3] DHH = Directorate of History and Heritage. CDI = Chief of Defence Intelligence. OSINT = open source intelligence. DSCA = Defense Security Cooperation Agency.

[4] Depending on whether a vessel that was bought for its parts was included.

Finally, in terms of depreciation, it would be useful to further explore the separation of the component that is due to aging design and technology relative to evolving functional requirements. This would be particularly useful if such modelling was of sufficient sophistication to account for the propensity of militaries to prolong the useful life cycle of platforms through betterments.

Beyond depreciated purchase cost, the entire life cycle costing of second hand platforms can potentially borrow techniques from DMGOR[5]'s stochastic models for operations and maintenance and prediction of optimal fleet renewal intervals. The disruptive transition between ownership periods and its effect on the modelling approaches warrants further study.

As mentioned in the Background, methods developed and used for specific platforms can be adapted and extended for use on other platforms. It can also help determine an initial price range if the CAF was to consider selling assets.

[5] DMGOR = Directorate Materiel Group Operational Research.

Sommaire

Sale price as a component of life cycle costing for second hand platforms : Depreciation modelling in the context of the FFG-7 frigate

Fred Ma ; DRDC CORA TM 2013–227 ; R & D pour la défense Canada – CARO ; décembre 2013.

Contexte : En mai 2012, le Directeur - Politiques et procédures (Matériel) a entrepris une étude de comparaison des achats d'équipement majeur neuf aux achats d'équipement usagé. L'un des principaux éléments d'une telle étude est l'analyse des coûts. On peut appliquer une méthode quantitative d'analyse des coûts à l'achat de diverses plateformes, par exemple des frégates, des navires de soutien, des destroyers, des véhicules et des aéronefs. Une telle analyse financière est essentielle lorsqu'on veut faire une comparaison complète tenant compte de facteurs supplémentaires comme la pertinence opérationnelle d'une plateforme candidate de seconde main (c.-à-d., offre-t-elle les capacités requises, maîtrise-t-elle suffisamment ces capacités et dans quels délais la capacité opérationnelle totale peut-elle être atteinte[6]).

Objectifs : Le présent rapport offre une description de l'étape préliminaire de planification de la modélisation des coûts du cycle de vie lorsqu'on achète une plateforme majeure de seconde main. On y présente aussi une étude de cas portant sur la prévision d'un élément de coût majeur, soit le prix de vente d'une plateforme, en l'occurrence la frégate FFG-7. Dans le présent rapport, on considère la dépréciation de la plateforme comme un synonyme du prix de vente d'équipement de seconde main[7].

Méthode : Pendant la planification préliminaire, on a dégagé les facteurs pouvant influencer le coût du cycle de vie d'une plateforme de seconde main. À partir d'un sous-ensemble de ces facteurs, on a établi une petite base de données permettant d'analyser la dépréciation en tant que conséquence de l'âge de la plateforme. Les principaux facteurs sont l'âge (dates d'achat à l'état neuf ou à l'état usagé) et le coût (neuf et usagé). Au départ, on avait désigné deux plateformes navales pour lesquelles on croyait être en mesure d'obtenir suffisamment de données, à savoir la Kortenaer des Pays Bas et la FFG-7 des États Unis. On a d'abord effectué une recherche de données dans des sources ouvertes afin d'alimenter la base de données (seuls les coûts d'achat à l'état neuf et à l'état usagé de la FFG-7 étaient faciles à trouver). Par la suite, on a employé un modèle simple de décroissance exponentielle pour établir la dépréciation en dollars constants.

Limites de la méthode : L'une des grandes difficultés de cette étude était de trouver les coûts d'achat à l'état neuf et à l'état usagé des plateformes de seconde main. Par ailleurs, les plateformes de

[6] Cela comprend les restaurations prévues, la canadianisation (ou, en d'autres termes, les modifications à apporter en vue de combler les exigences particulières du pays acheteur) et les délais auxquels on peut s'attendre, selon les précédents historiques, dans l'éventualité de travaux de restauration ou d'adaptation imprévus. Le présent rapport contient des exemples de cas montrant l'importance de ces considérations.

[7] On associe la dépréciation à la vente d'équipement usagé. Dans le rapport, on établit une distinction entre la dépréciation et la diminution générale de la valeur au fil du temps d'une plateforme particulière.

seconde main en question ont été achetées il y a longtemps, alors que les Forces armées canadiennes (FAC) chercheront plutôt à obtenir de l'information sur des plateformes qu'elles envisagent d'acheter. Il est possible que l'on ne retrouve pas beaucoup d'occurrences comparables dans le passé, ce qui compliquerait la cueillette de données.

La modélisation de la dépréciation touche à un seul élément du coût du cycle de vie, lequel représente en soi seulement un des aspects d'une évaluation des options d'achat, comme cela a été mentionné plus haut. En outre, l'analyse dont il est ici question aborde seulement un des facteurs de dépréciation (l'âge), bien que ce soit le principal. Enfin, les achats de plateformes usagées représentent seulement une fraction des achats de plateformes. En connaissant les caractéristiques générales de la détérioration de la plateforme dans le temps, on serait en mesure de brosser un tableau plus complet de l'achat d'équipement usagé, notamment une meilleure connaissance des risques et une meilleure idée de l'ampleur des recherches nécessaires sur l'état de la plateforme. On ne peut d'ailleurs pas dissocier la question des meilleurs indicateurs à employer pour mesurer la détérioration générale de celle de la disponibilité des données.

Résultats : On a créé une base de données pour les deux classes de frégates à partir de renseignements provenant de sources ouvertes et à l'aide du DHP, des responsables de l'OSINT au CRD et du DSCA, une agence américaine[8]. Comme il a été mentionné précédemment, il a été possible de trouver seulement les coûts d'achat neuf et usagé du FFG-7. Après un rajustement en fonction de l'inflation, on a modélisé la dépréciation réelle selon une courbe de décroissance exponentielle sur le plan du coût de vente de seconde main pour arriver à 8.4% par année, avec 68% d'intervalle de confiance de [5.9%, 11.1%] par année. On a tenté d'estimer la partie de la dépréciation attribuable à l'âge du modèle et de la technologie de la plateforme par rapport à l'évolution des besoins militaires, mais l'échantillon était trop petit et s'échelonnait sur une période trop courte. Cependant, on n'a trouvé aucune donnée réfutant la croissance annuelle de 2% du coût réel des classes de navires visées ayant été achetées neuves.

Les développements théoriques dans la validation du modèle sont présentés dans les annexes.

Portée : La croissance annuelle de 2% du coût des navires achetés neufs figure parmi les plus bas taux de croissance de coût dans le domaine de l'équipement de défense. En effet, il n'est pas rare de voir des taux de croissance annuels de 5 à 6% ; pour certaines classes d'équipement, le taux peut dépasser les 10%. C'est pourquoi il est d'autant plus important d'analyser le plus minutieusement possible les différentes options d'achat pour maximiser la fiabilité des prévisions des coûts du cycle de vie, y compris pour les achats d'équipement de seconde main. Grâce à un modèle simple de décroissance exponentielle, il a été possible d'expliquer 41% à 56%[9] du prix de vente d'équipement usagé à l'aide d'une seule variable d'entrée (l'âge du navire).

Travaux futurs : Plus on a de données, plus la modélisation sera fiable. Ainsi, il n'y a pas que l'âge de la plateforme qui peut servir à prévoir les coûts d'achat usagé. En effet, l'utilisation de différentes variables influençant potentiellement la dépréciation permet d'estimer le coût d'achat de plateformes

[8] DHP = Direction - Histoire et patrimoine ; CRD = Chef du renseignement de la Défense ; OSINT = renseignement de sources ouvertes ; DSCA = Defense Security Cooperation Agency.

[9] Selon qu'il s'agisse d'un navire acheté pour ses pièces ou non.

usagées autres que celle pour laquelle on a amassé des données. Pour une modélisation d'une telle ampleur, il faut plus de données, mais cette approche a été éprouvée par l'estimation du coût d'achat de nouvel équipement (le navire de soutien interarmées).

Pour améliorer l'estimation du prix de vente d'équipement de seconde main, il serait utile d'établir des approches permettant de tenir compte des accroissements de la valeur et de l'équipement complémentaire accompagnant la plateforme usagée, des éléments qui font augmenter l'écart des estimations de la dépréciation en fonction du prix de vente à l'état usagé. Il peut être possible de représenter cette dépendance en ajoutant à un modèle multidimensionnel général des variables d'entrée judicieusement définies.

Enfin, pour ce qui est de la dépréciation, il serait utile d'explorer plus en profondeur la séparation de l'élément découlant de l'âge du modèle et de la technologie par rapport à l'évolution des besoins fonctionnels. Cela serait particulièrement utile de disposer d'une méthode de modélisation assez sophistiquée pour tenir compte de la propension des militaires à prolonger le cycle de vie utile des plateformes au moyen d'améliorations.

Outre la dépréciation du coût d'achat, on pourrait établir le coût du cycle de vie complet d'une plateforme de seconde main en empruntant des techniques des modèles stochastiques du DROGM[10] pour le fonctionnement et l'entretien et la prédiction des intervalles optimaux de renouvellement des flottes. La transition désorganisée entre les propriétaires et ses effets sur les méthodes de modélisation mérite également qu'on y consacre une étude.

Comme cela a été mentionné dans la partie « contexte », les méthodes utilisées pour des plateformes particulières peuvent être adaptées pour d'autres plateformes. Elles peuvent également servir à déterminer une fourchette de prix initiale dans l'éventualité où les FAC envisageaient de vendre de l'équipement.

[10] DROGM = Directeur - Recherche opérationnelle (Groupe des matériels).

Table of contents

List of tables

List of figures

Acknowledgements

The author is indebted to the following people:

- Ms. Manchun Fang for discussions on multicollinearity in multiple regression;
- Mr. Brian Carter (DMPP SAPP) for background on procurement of major military platforms and costing;
- Mr. David Shaw (DMGOR 2) for his extensive guidance and assistance in terms of both analysis, military knowledge, DMGOR reporting standards, and transformative feedback on the first draft of this report;
- Dr. David Maybury (DET) and Mr. David Shaw for their guidance on currency equivalences across time and between nations, and for extensive discussions on modelling with stochastic differential equations; and
- Dr. Paul E. Desmier (DMGOR) for connecting the author to this project, providing all the background papers, for the omnibus normality testing, and for transformative feedback on this report.

This page intentionally left blank.

1 Introduction

In May 2012, Director Materiel Policy and Procedures (DMPP) initiated a study into the trade-offs of purchasing new versus used major equipment [1]. High level factors being considered in the study include purchase cost, timeliness of achieving full operational capability (FOC), and how well the platform capabilities fit CAF operational requirements[11]. This report describes an analysis of one contributor to the cost of buying used, namely sale price as a function of age. Due to the availability of data, the method was applied to only one platform, the U.S.'s Oliver Hazard Perry-class frigate (FFG-7) [2]. The method is not platform-specific, and therefore can be applied to any type of platform. More sophisticated methods developed within DMGOR to model purchase price and life cycle costs of platforms bought new can potentially be adapted to analyze second hand purchases. A prerequisite for these analysis methods is data. In particular, new and used purchase price data seems to be challenging to obtain.

1.1 Background: The larger plan

A larger cost modelling plan, of which this analysis was a part, has been put on hold due to the lack of open source data on new and used purchase prices. It is described here for context, and in case such a study is undertaken in the future.

In the preliminary planning of a model for life cycle costing of second hand platforms, a particular pattern of costs was anticipated (Figure 1). The initial purchase cost of buying used can be much lower than buying new. However, additional costs and delay can be expected in the initial years to restore the platform to operational condition and for required domestic adaptions, betterments, testing, and accreditation. Cases examined in the course of study [1] suggest that unanticipated costs and delays for adaptions can be significant, and there is also significant risk that operational requirements will not be met[12]. Beyond these up-front costs, the used platform is expected to have higher Operations and Maintenance (O&M) costs than a new platform for the remainder of its second life, and the life cycle will be shorter. The intent of the larger analysis plan was to develop this template of costs into a model, stochastic or deterministic, for exploring the life cycle cost of buying used compared to buying new. In discussions within DMGOR, it was acknowledged that costs are one aspect of a complete comparison that ideally includes factors that were difficult to predict: operational suitability, proficiency, and timeliness of FOC.

Two possible approaches to estimating contributing costs of buying used are the data mining methods of reference [5, 6] for shipbuilding, and the optimal life cycle estimation of references [7, 8] for light armoured vehicles and maritime surveillance aircraft, suitably adapted for used assets. Sales information was expected to be most readily available for the Dutch Kortenaer-class frigate [9] and the FFG-7. Hence, data gathering was undertaken for these two ships for this analysis.

[11] Described in [1] as cost, schedule, and capability.

[12] A domestic example is the prolonged service of aging Leopard 1 tanks due to incompatibility between Leopard 2's and the CAF's mine ploughs, bulldozer blades, and mine rollers [3]. An Australian example is the conversion of a commercial tanker into an Auxiliary Oiler [4], ahead of schedule and under budget but without the aviation support, fuel compatibility, speed, and manoeuvrability to fulfil its role in a task group. It will be replaced with a purpose built ship, but in the interim, the service of a operationally suitable ship has to be extended well beyond its retirement date.

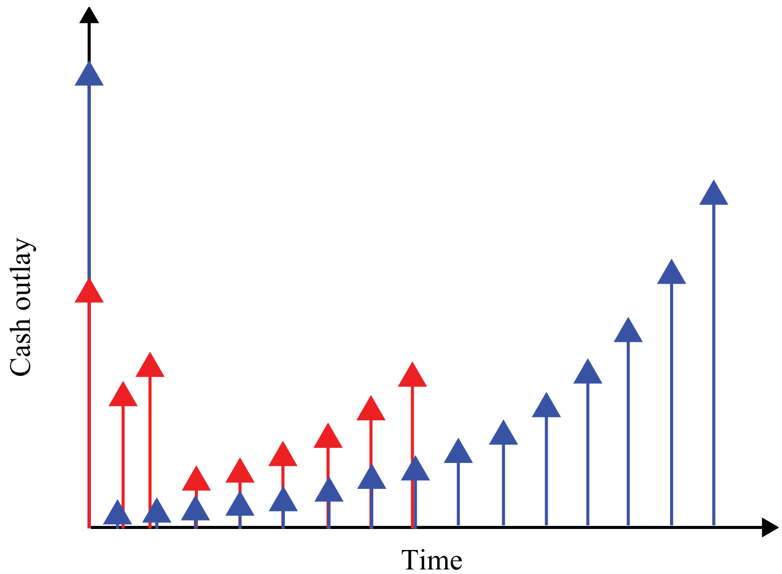

Figure 1: *Notional life cycle cost profiles for buying used (red) versus new (blue).*

In references [5, 6], data mining algorithms were used to estimate the ship development and construction costs for the Royal Netherlands Navy Rotterdam class and for the notional CAF Joint Support Ship (JSS). The estimates were based on multinational data sets of dozens of somewhat similar ships. In terms of an abstracted problem, the single output parameter of the model was the forecasted cost of a new ship. The forecast was made based on the dependence of the cost on a wide variety of input parameters, both technical and physical. If instead the output is taken to be *used* ship cost, a number of additional input parameters could be included for its prediction. Age of the platform at the time of sale and price-when-new were the most obvious additional input parameters, but so were selling country, buying country, year of sale, whether the sale was intra-country, the shipyard, and the rank in class[13]. While these additional fields were considered in planning a database for the input data, it would be simple to disregard any of them in the analysis if (i) they did not explain the variance in the output enough to warrant the extra dimensionality or (ii) there was not enough data with which to populate the higher dimensional model.

In references [7, 8], annual O&M costs for vehicles and aircraft were treated as stochastic functions of age. The variation in annual O&M costs was treated as geometric Brownian motion to account for the random nature of maintenance requirements. A number of potential challenges are identified below in the use of this method to estimate the life cycle cost of buying used. These challenges, and solutions to them, are the potential subjects of future studies into adapting stochastic models to second hand purchases.

1. **Obtaining data:** The stochastic model depends on the availability of historical O&M cost data in order to extract the model parameters. This data corresponds to the blue cost profile

[13] Experience with a ship class accumulates with ships built, making later ships less costly than initial ships. This has been studied as a learning / cost improvement effect for aircraft [10, 11], and as a dependence on rank-in-class for ships [5, 6].

in Figure 1 up to the year in which the first owner ceases to operate the platform and sells it. The red profile represents the costs for the buying nation for the second hand purchase and subsequent adaptions and O&M costs.

2. **Disruptive suspension of service:** The decommissioning and recommissioning of major equipment disrupts trends in depreciation. Platforms in continuous operation and subject to regular inspection and maintenance degrade differently (not necessarily more) than platforms that are simply stored for possible resale. Hence, the duration of discontinued operation could serve as a useful input variable in the above data mining approach to forecasting the used cost.

3. **Different incentives for O&M between owners:** Over the first life of major assets, original owners who intend to sell after a limited period of operation may not maintain the platform in the same manner as owners intending to optimize the cost of ownership over a longer life cycle. This can be likened to the "moral hazard" of economic theory, wherein parties tend to behave less responsibly if they do not personally bear the consequences [12].

4. **Memorylessness:** Due to the memoryless nature of Brownian motion, it may difficulty to model gross temporal patterns in spending that are highly plausible in a resale situation. Brownian motion consists of a sequence of random permanent level shifts, each of which are independent and identically distributed [13]. Applied to a platform fleet's annual O&M costs, if the cost is low in one of its later years, the future O&M costs are probabilistically the same as that of a much younger fleet[14]. This could be problematic if there is a tendency for first owners to curb maintenance in the final year(s) of operation, thus introducing a bias and anomaly for the final data point from the first period of ownership. Whether this can be taken into account in the forecasting of costs for the second period of ownership was not clear at the time of writing, especially considering that more sophisticated models typically require more data.

5. **CAF adaptions:** In contrast to new acquisitions, making used platforms operational may require significant adaptions (anticipated or not) to meet CAF requirements.

1.2 Used purchase price versus age: The near term plan

The immediate step in the larger study was to determine trends in the used purchase price versus age for the Kortenaers and FFG-7s. The determination of a trend should not be reduced to the development of a depreciation formula for accounting purposes because the resulting figures would not necessarily reflect actual operational value to the purchaser. As a reflection of operational value, used purchase prices were extremely difficult to come by in the open literature. Most of the new and used cost data that was gathered was for the FFG-7s, so only they could be analyzed for a depreciation trend.

According to the open source [2] (Wikipedia), the U.S. built the majority of the existing FFG-7s, whereas other ships based on the FFG-7 design were built by Australia; Taiwan, and Spain. However, from the open source (described next chapter, but primarily based on reference [2]), only the U.S.

[14] Such a process is known as a Martingale process [14]. At any instance t, the probabilistic future behaviour depends only on the value at t, regardless of what the historical values of the process were prior to t.

has sold FFG-7s used; the other nations have kept all theirs in service. Table 1 categorizes the fates of the 55 U.S. FFG-7s. The 17 that were transferred to other nations are the ones analyzed here. Three more transfers appear to be imminent; two of the frigates, one decommissioned and the other soon to be, have been granted to Mexico (transfer pending), and another decommissioned frigate is slated for foreign military sale.

Table 1: *Categorized fates of the 55 U.S. FFG-7s as per the open source [2] (Wikipedia). Cells specifying fates are highlighted as a visual aid. The yellow cells for the McClusky indicate a ship in transition.*

Name	Pennant	Fate	In service	Trans-ferred	Decom-missioned	Disposed of
Oliver Hazard Perry	FFG-7	Disposed of by scrapping, dismantling, 21 April 2006				✓
McInerney	FFG-8	Transferred to Pakistan as PNS Alamgir (F-260)		Pakistan		
Wadsworth	FFG-9	Transferred to Poland as ORP Gen. T. Kościuszko (273)		Poland		
Duncan	FFG-10	Transferred to Turkey as a parts hulk		Turkey		
Clark	FFG-11	Transferred to Poland as ORP Gen. K. Pułaski (272)		Poland		
George Philip	FFG-12	Stricken, to be disposed of, 24 May 2004.				✓
Samuel Eliot Morison	FFG-13	Transferred to Turkey as TCG Gokova (F 496)		Turkey		
Sides	FFG-14	Stricken, to be disposed of, 24 May 2004.				✓
Estocin	FFG-15	Transferred to Turkey as TCG Goksu (F 497)		Turkey		
Clifton Sprague	FFG-16	Transferred to Turkey as TCG Gaziantep (F 490)		Turkey		
built for Australia as HMAS Adelaide	FFG-17	Decommissioned, sunk as diving & fishing reef, April 2011			✓	
built for Australia as HMAS Canberra	FFG-18	Decommissioned, sunk as diving & fishing reef, October 2009			✓	
John A. Moore	FFG-19	Transferred to Turkey as TCG Gediz (F 495)		Turkey		
Antrim	FFG-20	Transferred to Turkey as TCG Giresun (F 491)		Turkey		
Flatley	FFG-21	Transferred to Turkey as TCG Gemlik (F 492))		Turkey		
Fahrion	FFG-22	Transferred to Egypt as Sharm El-Sheik (F 901)		Egypt		
Lewis B. Puller	FFG-23	Transferred to Egypt as Toushka (F 906)		Egypt		
Jack Williams	FFG-24	Transferred to Bahrain as RBNS Sabha (FFG-90)		Bahrain		
Copeland	FFG-25	Transferred to Egypt as Mubarak (F 911), renamed Alexandria in 2011		Egypt		
Gallery	FFG-26	Transferred to Egypt as Taba (F 916)		Egypt		
Mahlon S. Tisdale	FFG-27	Transferred to Turkey as TCG Gokceada (F 494)		Turkey		
Boone	FFG-28	Decommissioned 23 February 2012			✓	
Stephen W. Groves	FFG-29	Decommissioned 24 February 2012			✓	
Reid	FFG-30	Transferred to Turkey as TCG Gelibolu (F 493)		Turkey		
Stark	FFG-31	Disposed of by scrapping, dismantling, 21 June 2006				✓
John L. Hall	FFG-32	Decommissioned 9 March 2012			✓	
Jarrett	FFG-33	Decommissioned, held for future foreign military sale			✓	
Aubrey Fitch	FFG-34	Disposed of by scrapping, dismantling, 19 May 2005				✓
built for Australia as HMAS Sydney	FFG-35	In active service, as of 2013	✓			
Underwood	FFG-36	Decommissioned Mar 8, 2013			✓	
Crommelin	FFG-37	Decommissioned October 26, 2012			✓	
Curts	FFG-38	Decommissioned January 25, 2013. Granted to Mexico in 2013 but transfer pending.			✓	
Doyle	FFG-39	Decommissioned July 29, 2011			✓	
Halyburton	FFG-40	In active service, as of 2013	✓			
McClusky	FFG-41	Ship in active service. Set to be decommissioned in 2014. Granted to Mexico for 2014 but transfer pending.	✓			
Klakring	FFG-42	Decommissioned Mar 22, 2013			✓	
Thach	FFG-43	In active service, as of 2013	✓			
built for Australia as HMAS Darwin	FFG-44	In active service, as of 2013	✓			
De Wert	FFG-45	In active service, as of 2013	✓			
Rentz	FFG-46	In active service, as of 2013	✓			
Nicholas	FFG-47	In active service, as of 2013	✓			
Vandegrift	FFG-48	In active service, as of 2013	✓			
Robert G. Bradley	FFG-49	In active service, as of 2013	✓			
Taylor	FFG-50	In active service, as of 2013	✓			
Gary	FFG-51	In active service, as of 2013	✓			
Carr	FFG-52	Decommissioned Mar 13, 2013			✓	
Hawes	FFG-53	Decommissioned, to be cannibalised in Philadelphia			✓	
Ford	FFG-54	In active service, as of 2013	✓			
Elrod	FFG-55	In active service, as of 2013	✓			
Simpson	FFG-56	In active service, as of 2013	✓			
Reuben James	FFG-57	To be decommissioned Aug 30, 2013			✓	
Samuel B. Roberts	FFG-58	In active service, as of 2013	✓			
Kauffman	FFG-59	In active service, as of 2013	✓			
Rodney M. Davis	FFG-60	In active service, as of 2013	✓			
Ingraham	FFG-61	In active service, as of 2013	✓			
			19	17	14	5
				55		

2 Data gathering

Acquiring data from open sources was a major challenge for used ship sales, and this was particularly true for the new and used purchase prices. An Excel database was planned for used sales of both Kortenaers and FFG-7s. New and used cost information specifically was unavailable for most of the Kortenaers. The fields for the database are shown in Table 2.

Table 2: *Data collection plan for used sales.*

Data about:	Fields
Platform	Platform (only frigates)
Providing nation	Nation Class Platform name Pennant
Receiving nation	Nation Class Platform name Pennant
Start of first service	Nation of construction Launch date Date of completion First ever commission date
Transition to most recent service	Preceding de-commission date Date of sale Handover date Recipient launch date Recipient commission date
Value [USD]	Cost new Used purchase price Assessed used value

There are a number of date fields associated with the start of first life and start of second life because of the various stages in operationalizing a vessel. These dates helped in estimating the start of operations if the exact date was not found, and they helped reconcile information from different sources. It was not uncommon for information to be incomplete or inconsistent between sources. One field, "Date of completion" was not precisely defined; hence it is simply a rough indicator that operations started some time around the year specified by that source. Similar judgement calls were sometimes made in determining which ship in its first life became which ship in its second life[15]. These judgement calls were informed by the weight of the evidence, e.g., the amount of the information from various sources and their credibility, both subjectively assessed and as advised by Chief of Defence Intelligence's (CDI's) Open Source Intelligence (OSINT).

The open source nature of the data meant that the database was built up in a very iterative manner. Whenever new information was encountered, whether it consisted of a few ships (e.g., press articles) or an entire database of ships on the web, each affected field of each affected record was re-assessed based on all the information accumulated for that cell. Open source information can be ephemeral to varying degrees, as can DND's access to even established sources; therefore, every effort was made to locally archive the source information in a repository, thus providing an audit trail and ensuring reproducibility of the analysis in the future.

One particularly noteworthy source of uncertainty in the used price data was the value-enhancing accessories accompanying each vessel and enhancements applied to each vessel, e.g., weapons, sensors, aircraft, refurbishments. There was no information with which to separate this out. For a more in-depth study, this would raise the question of what should be considered part of a major platform. Sometimes, when more than one vessel was sold at a time, it was unclear what enhancements were associated with which vessel; the vessels may be of different ages and have different used purchase prices.

2.1 Sources of data

In addition to various sources searched by OSINT, major sources of information for the database are shown in Tables 3 and 4. Personnel from the following organizations contributed to the open source search:

1. Directorate of History and Heritage (DHH);
2. CDI's OSINT; and
3. Defense Security Cooperation Agency (DSCA).

DSCA's Excess Defense Articles (EDA) database was the source of all of the used purchase prices for the FFG-7s.

[15] The second owner typically assigns a different name and pennant. There was sometimes conflicting information about the mapping of individual ships in their first lives to individual ships in their second lives.

Table 3: *FFG-7 data sources.*

URL	Organization	Information
turkishNavy.net	No affiliation	Pennant, name, launch date, commission date, buying nation
wikipedia.org	Wikimedia Foundation	Pennant, name, various dates, fate, various information on ship-specific pages
www.dsca.mil/programs/eda/edamain.htm	Defense Security Cooperation Agency	Cost (new, used), date of used sale, news releases
www.globalsecurity.org	Independent organization dealing in information on defence and security	Articles
www.nvr.navy.mil	The Naval Vessel Register (NVR), official inventory of ships and service craft	Various dates, various technical info available

Table 4: *Kortenaer data sources.*

URL	Organization	
janes.ihs.com	IHS Jane's	Articles
wikipedia.org	Wikimedia Foundation	Name, pennant, various dates, fate
www.amiinter.com	AMI International (naval information products/services)	Google linked articles with cost
www.bicc.de	Bonn International Center for Conversion	Dutch surplus weapons document (various info on transfers)
www.hellas.org	No info on host organization Blocked as "Extreme, Politics/Opinion" by DRENET	Pennant, name, commission date
www.navyleague.org	The Navy League of the United States	Assorted articles with pricing info
www.seaforces.org	Seaforces online (independent information synthesis, various navies)	Name, possibly various dates
www.worldnavalships.com	No information on host organization	Name, various dates, pennant, fate

3 Inferring second hand cost versus age

3.1 Conversion of dollar values to a common year

The new and used ship costs were inflated to 1 Jan 2012 using the Consumer Price Index (CPI) inflation percentages in Annex A [15]. Note that InflationData.com, the source of this data, calculated inflation rates for a calendar year by arithmetically averaging the rates over the months of a given year rather than geometrically averaging them. This is because the rate for each month is not the inflation during the month; it is the inflation over one year between that month and the same month of the preceding year.

For this study, the calendar year inflation rates in Annex A were applied geometrically to inflate cash flows to 1 January 2012. The date of each cash flow was rounded to the nearest start of a calendar year. If payment occurred in the first 6 months of a year Y, it was treated as if it occurred at the start of the year; if payment was made in the last 6 months of a year Y, it was deemed to have happened at the start of year $Y + 1$.

At the time of writing, two new sources of inflation rates that were specific to shipbuilding were obtained from Kaluzny [16]. The first new source was the inflation index used in reference [6]: the Historical Shipbuilding and Conversion, Navy (SCN) Total Obligational Authority (TOA) Index [17], used within the U.S. Naval Sea Systems Command (NAVSEA). A quick check of this new data was done to confirm that these rates was in line with the ones from [15]. Figure 2 shows the correspondence between these two sets of inflations rates.

The second new source was a Joint (services) Inflation Calculator hosted by the Naval Center for Cost Analysis [18]. The knowledge required to use this calculator was explored, but the calculator was not pursued for this study. A significant amount of time, and likely some liaising with the host organization, would be needed to become familiar with the budgeting and cost categorizations, weighting schemes, and assumed temporal spending patterns in order to properly generate inflation data that was suitable for this study with any degree of confidence. Furthermore, much of the terminology and many of concepts parallelled those found in the NAVSEA data above.

Other inflation indices found online can be more suitable for ship building than the CPI. These include:

1. Shipbuilding Indices for steel vessel contracts [19], prepared by the U.S. Bureau of Labor Statistics (BLS) for NAVSEA
2. BLS's Producer Price Index (PPI) for newly constructed military self-propelled ships [20]

However, these sources only go back to 1988 and 2003, respectively. For the FFG-7 data gathered, inflation data was needed back to 1980.

Note that the greater suitability of ship building cost figures is an assumption whose validity may depend on the context of the analysis. For example, if the option of buying used is being financially assessed in order to decide on how best to use funds the from a budget meant exclusively for ship procurement, then shipbuilding indices are a sensible option. However, if the budget for purchasing

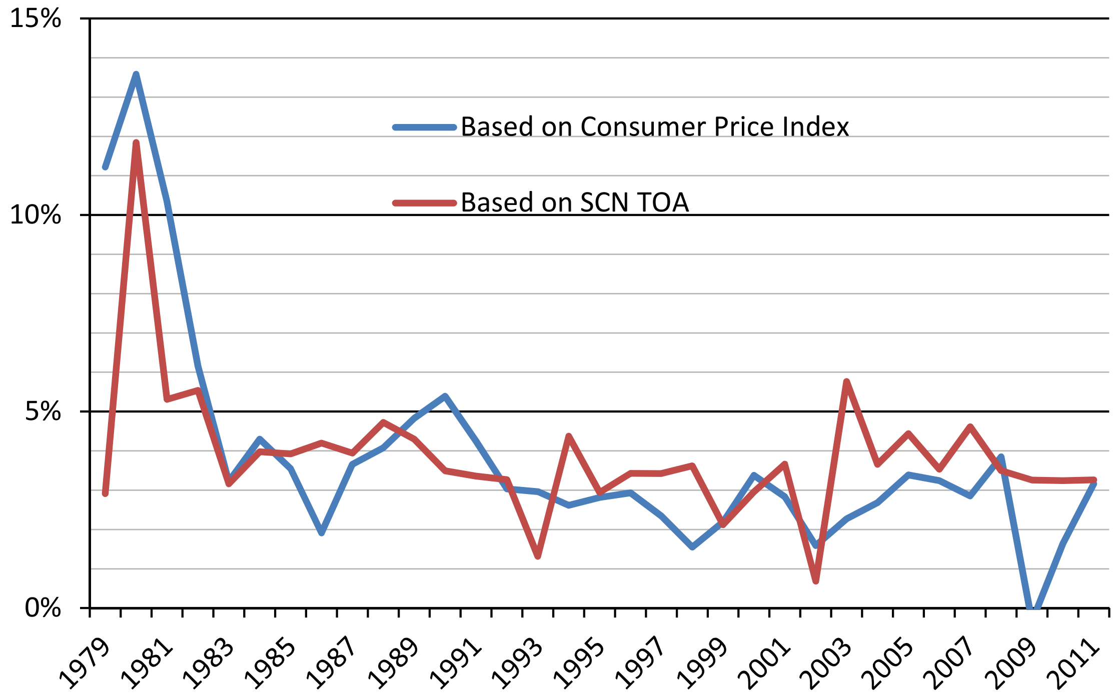

Figure 2: *Comparison of inflation rates from the Consumer Price Index and SCN TOA.*

second hand is among a set of budgets for various purposes, between which there is much leeway with which funds can be shifted, then the more general CPI may be more appropriate. In this particular study, the similarity between the indices makes the choice of little consequence. As a general observation, however, the analysis context will determine whether a sector specific index for an asset under study is more suitable than a general inflation index.

3.2 Analysis

In general, the simplest plausible depreciation model for a major asset is exponential decay relationship

$$Cost_{\text{Used}} \simeq a \cdot Cost_{\text{New}}\, e^{-b \cdot Age} \,, \tag{1}$$

where a and b are fitting constants that characterize the decay, and the approximate nature of the model is explicitly represented. Ideally $a \in [0, 1]$ and $b > 0$. The relative retained value at a given Age is then

$$\frac{Cost_{\text{Used}}}{Cost_{\text{New}}} \simeq a\, e^{-b \cdot Age} \; . \tag{2}$$

The relative retained value is just a complementary way of representing depreciation; whatever portion of the original value is not retained is the reduction in value, which is depreciation.

After determining a and b, a can be viewed as a drop in value if a ship was to be hypothetically re-sold immediately after being purchased ($Age = 0$), similar to the immediate drop in the value of a car the moment it is driven off a dealer's lot. However, if there is little or no data with Age near zero, then the fitted value for a might not satisfy $a \in [0, 1]$. For example, if $a > 0$, then unless there is a specific reason why an asset would appreciate immediately after being sold, a cannot be interpreted beyond being just a fitting parameter.

The component $e^{-b \cdot Age}$ is the Age dependent decay. Being the simplest possible model of decay, the constant value retained each elapsed year can be obtained by substituting $Age = 1$ year. The annual percentage retained value is then

$$\alpha \simeq 100\% e^{-b(1 \text{ year})} \; , \tag{3}$$

where the 100% factor merely scales the resulting fractional number so that it is expressed as a percentage, and time unit for the 1 year is explicitly retained as per standard practice in the physical sciences to ensure that b is properly scaled for Age expressed in years (it has units of 1/time). The annual depreciation is then

$$d_{1\text{yr}} = 100\% - \alpha \; . \tag{4}$$

In order to evaluate (3) and (4), we work with the log transformed version of (2),

$$\ln \frac{Cost_{\text{Used}}}{Cost_{\text{New}}} \simeq \ln a - b \cdot Age \; , \tag{5}$$

to find representative values for $\ln a$ and b using linear regression[16]. Table 5 shows the calculation applied to the FFG-7 ages and resale prices, with ship-specific Age expressed in terms of n years (a per-ship value). Under the *Effective dates & age* section, the *New* column refers to the date of the start of a ship's first life. It is determined from the data in the ship's *Start of first service* fields (see Table 2). Sometimes, judgement call is required, but for any given ship, the dates in these fields are always close to each other. Similarly, the date in the *Used sale* field is synthesized from a ship's *Transition to most recent service* fields. The *New* and *Used sale* dates determine the *Age n* at the time of the second hand sale. As mentioned, the *Budget year prices* (both new and used) in U.S. dollars are from DSCA's EDA database. The generation of the inflation columns are discussed in the

[16] Because the base e is ubiquitously used for exponential decay, we use base e for all exponentiation and logarithms in this report.

preceding subsection, and the 2012 USD prices ($Cost_{\text{New}}$ and $Cost_{\text{Used}}$) are obtained by applying the appropriate inflators to the corresponding budget year prices. These are then used to calculate the relative retained value (2) so that the log transformed regression (5) can be applied. The regression is shown in Figure 3.

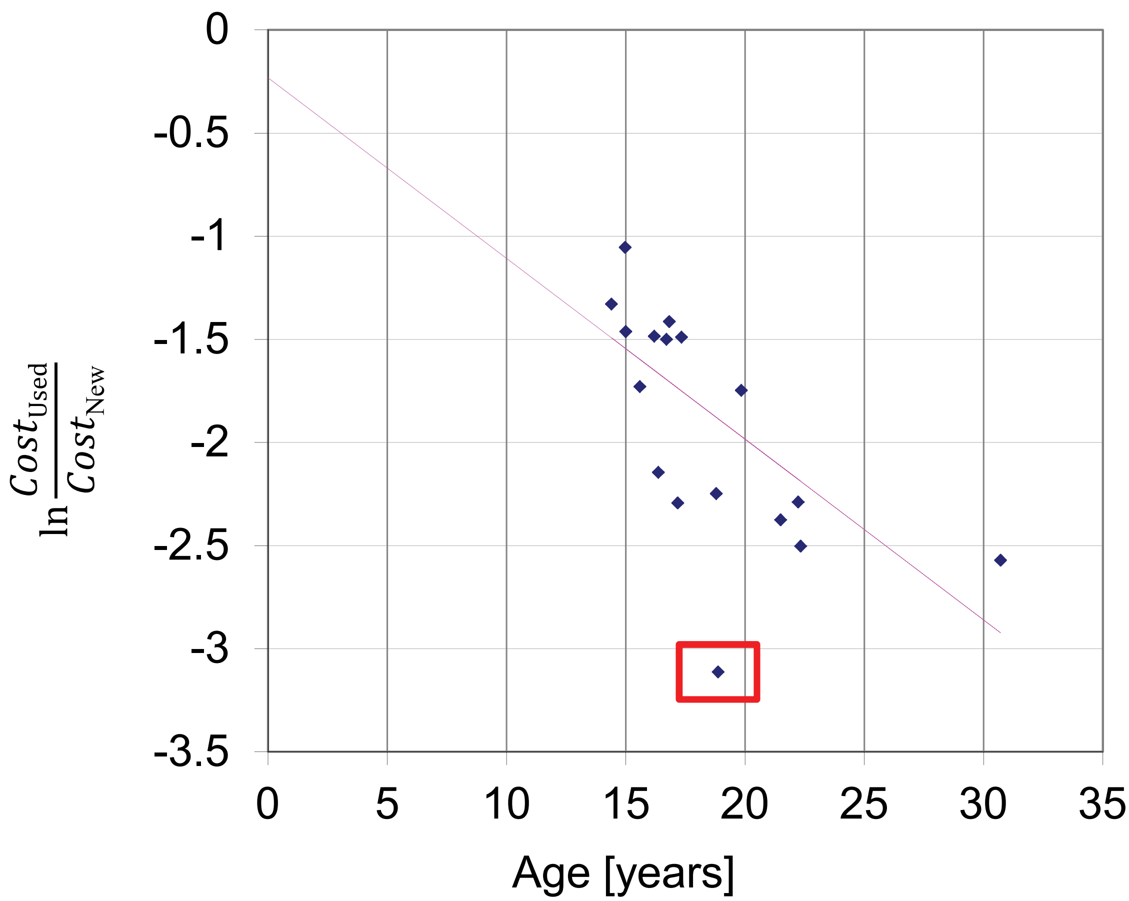

Figure 3: *Geometric depreciation of FFG-7 purchase prices. The highlighted point corresponds to an FFG-7 that was purchased as a parts hulk.*

One of the FFG-7s (highlighted) was purchased at an abnormally low second hand price. Subsequent investigation revealed that it was purchased for use as a parts hulk, i.e. it was purchased for its parts rather than for use as an operational vessel. Because of its anomalous fate, the analysis without the parts hulk is taken to be the default here, though results from the inclusion of the parts hulk are selectively discussed. We did not entirely dispense with the analyses that included the parts hulk because even if the intent is to purchase platforms for operational use, knowledge of how a particular platform depreciates in general can give an indication of the risk in buying used for that platform, especially if there is no guarantee that the buyer has detailed ground truth on the condition of the platforms.

The quantitative results of the regression are summarized in Table 6 for the cases with and without

Table 5: *Calculation of FFG-7 depreciation. The Age of each ship at the time of its used sale is designated n. The parts hulk is highlighted. Values in the depreciation columns are based on ship-specific prices (new and used), not the line of best fit.*

Pennant	Effective dates & age			Budget year price [USD]		% US inflation		Prices [2012-01-01 USD]		Annual real depreciation		
	New	Use sale	$Age\ n$ [years]	New	Used	Date new to 2012-01-01	Date of used sale to 2012-01-01	$Cost_{\text{New}}$	$Cost_{\text{Used}}$ at $Age\ n$	$\frac{Cost_{\text{Used}}}{Cost_{\text{New}}}$	α = % value retained / 1 year	1-α = % depreciation / 1 year
FFG-24	19/09/1981	13/09/1996	15.0	106,151,000	63,882,000	148%	43%	262,934,905	91,661,695	34.86%	93%	7%
FFG-22	16/01/1982	31/03/1998	16.2	134,057,000	53,622,800	148%	40%	332,057,772	75,181,927	22.64%	91%	9%
FFG-23	17/04/1982	01/01/1999	16.7	134,410,000	53,764,000	148%	38%	332,932,149	74,229,342	22.30%	91%	9%
FFG-25	07/08/1982	01/01/1997	14.4	136,961,000	58,929,000	133%	43%	319,565,702	84,554,836	26.46%	91%	9%
FFG-26	01/01/1982	01/01/1997	15.0	105,072,000	42,028,000	148%	43%	260,262,233	60,304,275	23.17%	91%	9%
FFG-8	15/12/1979	31/08/2010	30.7	115,469,000	26,568,943	210%	3%	358,479,033	27,408,522	7.65%	92%	8%
FFG-11	09/05/1980	15/03/2000	19.8	101,905,000	40,762,200	210%	35%	316,368,947	55,072,305	17.41%	92%	8%
FFG-9	28/02/1980	28/06/2002	22.3	136,536,000	27,307,200	210%	27%	423,882,543	34,705,352	8.19%	89%	11%
FFG-16	21/03/1981	24/07/1998	17.3	104,663,000	46,009,000	173%	40%	286,081,431	64,506,988	22.55%	92%	8%
FFG-20	26/09/1981	24/07/1998	16.8	140,541,000	60,361,000	148%	40%	348,118,571	84,629,231	24.31%	92%	8%
FFG-21	20/06/1981	27/08/1998	17.2	107,331,000	21,446,200	173%	38%	293,374,030	29,609,726	10.09%	88%	12%
FFG-30	19/02/1983	25/09/1998	15.6	149,435,000	44,830,500	133%	38%	348,670,795	61,895,293	17.75%	90%	10%
FFG-27	20/11/1982	05/04/1999	16.4	149,343,000	29,544,600	133%	38%	348,456,135	40,790,794	11.71%	88%	12%
FFG-19	14/11/1981	01/09/2000	18.8	135,283,000	27,056,600	148%	31%	335,094,561	35,360,006	10.55%	89%	11%
FFG-13	10/10/1980	11/04/2002	21.5	105,806,000	21,161,200	173%	27%	289,205,659	26,894,258	9.30%	90%	10%
FFG-15	10/01/1981	03/04/2003	22.2	94,930,000	21,021,143	173%	25%	259,477,659	26,298,116	10.14%	90%	10%
FFG-10	20/05/1980	05/04/1999	18.9	149,343,000	14,934,300	210%	38%	463,642,487	20,619,062	4.45%	85%	15%

the parts hulk. The y-intercepts have high p-values, so there is no evidence against the null hypothesis that the intercepts are zero. Since the parts hulk is the only outlier, the annual depreciation of 8.4%/year is not much affected by its inclusion. Of the FFG-7s sold for operational use, however, only 56% of the variance in depreciation is explained by the age of the ship (41% if the parts hulk is included). Hence, age alone is a very rough determinant of depreciation. The remainder of the variance in the log transformed data shows up in the standard deviation of the residuals. Normally, this indicates how far the actual depreciation is likely to be from fitted line. Since this is in the log domain, however, the standard deviation represents multiplicative factors after an antilog transform, as shown in Table 6 [17]. As expected, there is a greater spread in the case that includes the parts hulk.

Table 6: *Results of regression for FFG-7 depreciation.*

	With parts hulk	Without parts hulk
y-intercept) in log domain	-0.273	-0.231
P-value	0.61	0.56
Standard error in log domain	0.524	0.391
Variance in the log domain explained by age (r^2)	41%	56%
P-value	< 0.01	<< 0.01
Standard deviation of residuals[a] in log domain	0.444	0.331
Multiplicative effect on estimated depreciation	[0.6, 1.6]	[0.7, 1.4]
Estimated depreciation / year (from slope)	8.5%	8.4%
Standard error in log domain	0.028	0.021
68% confidence interval	[5.9, 11.1]%	[6.4, 10.3]%

[a] $\sqrt{(\textit{sum of squared residuals})/(N_{\text{Points}} - 2)}$.

As shown in Table 6, the limited explanatory power of the *Age* variable is also reflected in the broad 68% confidence intervals for both the initial loss and annual depreciation[18]. Here, too, the confidence intervals are broader for the case that includes the parts hulk.

3.3 Normality test results

Linear regression is based on the assumption of normally distributed residuals. The results of three normality tests are presented here. The details of the first two are contained in Annex C. Due to the exceptional nature of the parts hulk, it was not included in the normality checks.

Figure 4 shows the normal probability plot for the residuals[19]. The closeness of the dots to the red straight line is an indicator of the degree of normality in the residuals. The dots are very linear,

[17] The multiplicative factors give an intuitive sense of how symmetric offsets in the log domain affect the estimated depreciation. However, actual confidence intervals for estimated depreciation are *Age*-specific and are not summary statistics [21].

[18] Formulas for the standard error of the estimated slope and intercept are given in Annex B. Because of the small sample size, confidence intervals are determined using the t-distribution. The confidence interval boundaries are in terms of $x = \ln(\textit{Relative retained value})$ and are transformed back into depreciation as $1 - e^x$.

[19] It is good practice to always graphically examine the data as a complement to quantitative threshold tests [22–25]

with a coefficient of determination $r^2 = 95\%$. The pair of black rectilinear bounding lines are 95% confidence bands which the red line must reside between. The dots exhibit none of the gross features that indicate skewness or excess kurtosis.

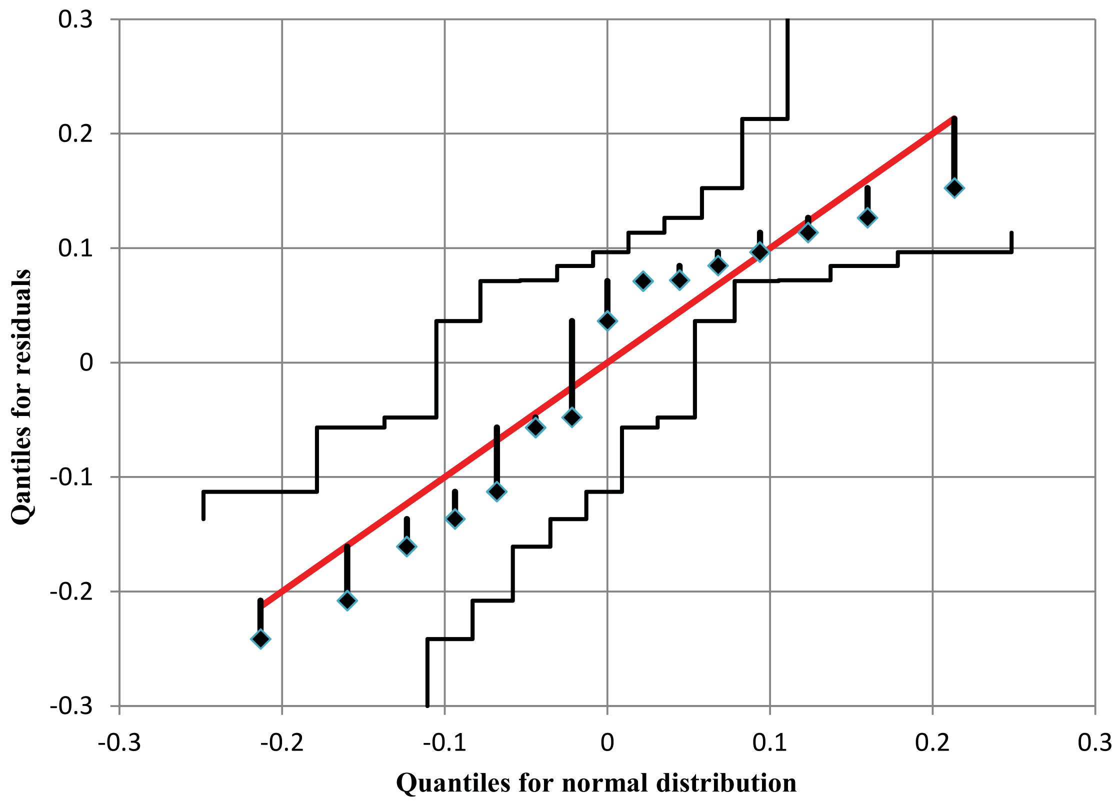

Figure 4: *Normal probability plot of residuals. The staircase shape of the 95% confidence bands arises because the discrete nature of the data points is prominent for small sample sizes.*

There is also no appreciable evidence against normality from the Anderson-Darling test, which works well for small sample sizes. The test statistic for our residuals exceeds none of the available thresholds for rejecting the null hypothesis of normally distributed residuals, up to a significance of 15%.

Finally, following reference [7], we also apply the recently developed Doornik-Hansen omnibus normality test [26] based on the 3rd and 4th moments of the data, skewness and kurtosis. The test is applicable to samples as small as 8. For our residuals, Table 7 contains the p-values from the test. The high p-values indicate that there is no appreciable evidence against normality nor against the absence of skewness and excess kurtosis.

Table 7: Goodness-of-fit metrics from omnibus test.

H_0	p-value
No skewness	0.834
No negative skewness	0.581
No positive skewness	0.419
No kurtosis	0.796
No negative kurtosis	0.602
No positive kurtosis	0.399
Data are normally distributed	0.947

4 Sensitivity analysis

This section describes the alternative analysis conducted to confirm that additional independent variables are not essential in the linear regression model, and that the different inflation data introduces a small amount of uncertainty. The analysis is performed on the default case, in which the parts hulk is excluded.

4.1 Multiple Regression

Apart from the age of the ship when it is sold second hand, the other fields in the database which depreciation could be regressed against are:

1. calendar year when a ship is new,
2. price when new,
3. year of second hand purchase,
4. second hand price, and
5. purchasing country.

Classical multiple regression finds the coefficients of dependence of the dependent variable, depreciation, on the above (ideally) independent variables such that the residuals are minimized. For a meaningful model, it is necessary to avoid multicollinearity [27, 28], where there is nontrivial correlation between variables. In an extreme illustration, imagine replicating a variable x and disguising it as a different variable y in the model. If the coefficient for x was (say) $b_x = 1$ in the absence of y, then introducing y into the model yields the same minimal residual regardless of whether coefficients were $\{k_x = 1, k_y = 0\}$, $\{k_x = 0, k_y = 1\}$, or $\{k_x = 0.5, k_y = 0.5\}$. In practice, x and y can be significantly correlated without being equal, yielding a subspace of solutions that are nearly optimum, where the coefficients have high variance, and the solver may have trouble accurately solving the ill-conditioned system. Because of possible correlation between variables selected as inputs to the model, they are less ambiguously referred to as predictor variables (rather than independent variables, which could be confused with the notion of statistical independence).

A number of steps were taken to vet and prepare the data for multiple regression. To avoid multicollinearity, we excluded the year of second hand purchase because it is highly correlated with age. We also excluded the second hand price as a predictor variable, not only because it is strongly correlated with depreciation, but also because that is what the model is intended to predict via the depreciation. The year when a ship is new was taken relative to the earliest year in the database: 15 Dec 1979, the date taken to be when the ship with pennant *FFG-8* was new. Finally, the purchasing country could not be used directly because it is a nominal variable, having no ratio or ordinal information. The typical way to incorporate such categorical data into regression is to define one binary variable per nation (known as a dummy or one-hot variable) with the exception of one nation, which is taken to be the nation that is implied when all of the binary variables are zero.

Table 8: *Correlation between predictor variables.*

Variable	ageYrs	relYrNew	usd2012new	Definition
ageYrs	1.00	-0.80	0.29	Age [years] when sold
relYrNew	-0.80	1.00	-0.10	Year when new, relative to 15 Dec 1979
usd2012new	0.29	-0.10	1.00	Price when new, inflated to 2012 and expressed as a factor of US$100,000,000.

The retained quantitative variables are shown in Table 8 along with their pairwise correlations. There is significant negative correlation between the year when a ship was new (*relYrNew*) and the age at which it was sold second hand (*ageYrs*), meaning that later ships were re-sold at an earlier age (Figure 5 left). There is also moderate correlation between the price when new (*usd2012new*) and *ageYrs* (Figure 5 right). We investigated the effect of including and excluding the highly correlated *relYrNew*.

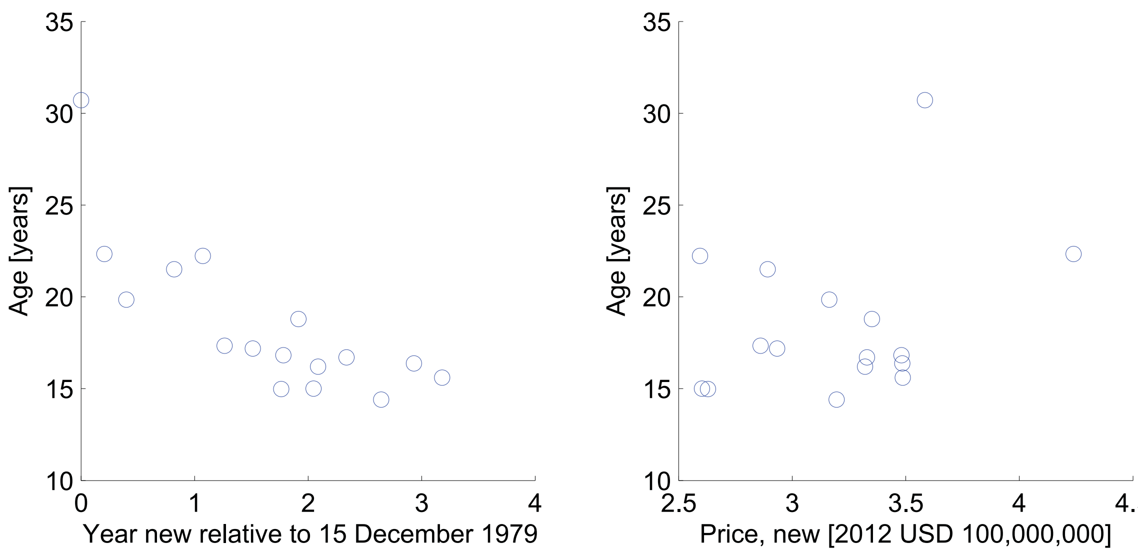

Figure 5: *Scatter plots of correlated pairs of variables.*

Table 9 shows the predictor variables investigated and their coefficients for six models. N/A means that the variable is not included in the model. Five of the variables are dummy variables corresponding to the purchasing nations, so only one such variable can be be non-zero, with that non-zero value being 1. In the case of Bahrain, N/A has meaning when the other nations are included in the model, which is the case in three of the models. In such models, Bahrain is the implied nation if none of the dummy variables for the other nations are non-zero. Also shown in Table 9 are figures of merits for each model, including the explanatory power R^2, the adjustment $\overline{R^2}$ of R^2 to account for the number of predictors used to achieve R^2, and the root of the mean of the squared residuals (the residuals are prediction errors and the radicand is designated *MSE* for mean squared error).

Table 9: *Summary of coefficients from regression models (one model per row). The p-values in brackets are rounded to 2 decimal digits.*

	ageYrs	relYrNew	usd2012new	Bahrain	Egypt	Pakistan	Poland	Turkey	Intercept	R^2	$\overline{R^2}$	$\sqrt{\text{MSE}}$
1	-0.09 (0.00)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	-0.23 (0.56)	0.56	0.53	0.33
2	-0.11 (0.03)	N/A	N/A	N/A	-0.33 (0.34)	0.16 (0.84)	-0.42 (0.36)	-0.60 (0.10)	0.55 (0.45)	0.76	0.65	0.29
3	-0.11 (0.01)	-0.11 (0.46)	N/A	N/A	N/A	N/A	N/A	N/A	0.34 (0.70)	0.58	0.52	0.34
4	-0.08 (0.00)	N/A	-0.00 (0.45)	N/A	N/A	N/A	N/A	N/A	0.20 (0.77)	0.58	0.52	0.34
5	-0.16 (0.04)	-0.19(0.41)	-0.06 (0.84)	N/A	-0.17 (0.65)	0.66 (0.49)	-0.34 (0.58)	-0.40 (0.34)	1.79 (0.19)	0.80	0.62	0.30
6	-0.12 (0.03)	N/A	-0.19 (0.42)	N/A	-0.23 (0.53)	0.50 (0.58)	-0.16 (0.78)	-0.47 (0.24)	1.19 (0.28)	0.78	0.63	0.29

Table 10: *Searches for good values for lasso's Elastic Net hyperparameters and the resulting coefficients.*

α	ageYrs	relYrNew	usd2012new	Bahrain	Egypt	Pakistan	Poland	Turkey	Intercept	λ	DF	MSE	SE	$\sqrt{\text{MSE}}$
						***relYrNew* excluded**								
0.01	-0.11	N/A	-0.18	N/A	-0.19	0.38	-0.18	-0.46	0.99	1.7×10^{-2}	6	0.13	0.032	0.36
0.33	-0.12	N/A	-0.19	N/A	-0.22	0.49	-0.15	-0.46	1.16	1.4×10^{-3}	6	0.13	0.032	0.36
0.67	-0.12	N/A	-0.19	N/A	-0.22	0.49	-0.15	-0.46	1.17	7.8×10^{-4}	6	0.13	0.032	0.36
1.00	-0.12	N/A	-0.19	N/A	-0.22	0.50	-0.15	-0.46	1.17	5.2×10^{-4}	6	0.13	0.032	0.36
						***relYrNew* included**								
0.01	-0.04	0.06	-0.17	N/A	0.08	-0.08	-0.07	-0.23	-0.47	0.44	7	0.14	0.034	0.37
0.33	-0.08	0	-0.14	N/A	0	0	-0.13	-0.33	0.18	0.06	4	0.15	0.036	0.38
0.67	-0.06	0	-0.04	N/A	0	0	0	-0.15	-0.49	0.13	3	0.15	0.037	0.38
1.00	-0.06	0	-0.01	N/A	0	0	0	-0.14	-0.55	0.10	3	0.15	0.037	0.38

In interpreting the coefficients, it rarely makes sense to interpret the magnitude of the coefficient in isolation, as that figure depends on the scaling units of the predictor variable. The p-value is the more reliable indicator of statistical significance. If we take $p = 0.1$ as the onset of evidence against a zero coefficient, then with one marginal exception discussed below, the age of a ship is the only variable that is significant. It is the only variable in all the models, and it is decidedly significant in all the models. The first model in Table 9 is the simple linear regression against the age of a ship in isolation, already presented. Each of the second through fourth models add a second variable in the following order: (i) purchasing nation, (ii) date when the ship was new (*relYrNew*), and price when new (*usd2012new*). The fifth model regresses depreciation against all the variables, while the sixth and final model uses all variables except for the highly correlated *relYrNew*.

The second model brings in the purchasing nations, which is contrasted against the first model in Figure 6. The actual data are shown as points, distinguished by symbology and colour according to nation, while the predictions are shown as solid lines. Predictions are made across the age range assuming that each nation in turn is purchasing. Since the each nation is a binary variable, the corresponding coefficient merely introduces nation-specific vertical offset into the linear dependence of depreciation on age. As expected, each line is centred through each nation's data points. We also note that the nation-specific slopes are steeper than the simple linear regression because nation-specific sub-sets of the data points are more steeply inclined.

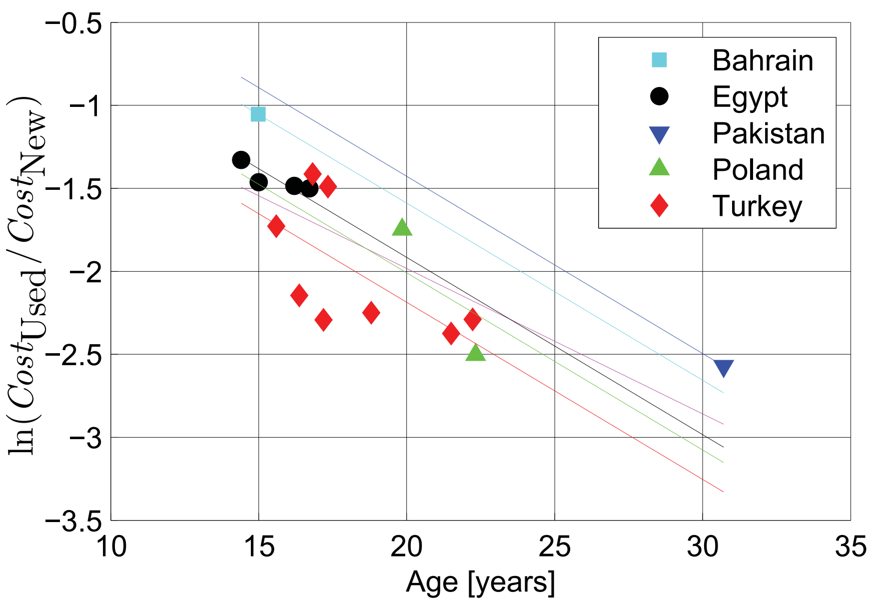

Figure 6: *Depreciation regressed against age of ship and purchasing nation. For comparison, the dashed line shows the simple linear regression against age alone.*

The nation-specific offsets are graphically satisfying, but with the exception of Turkey, the p-values indicate that the nation effects are well within sampling noise. However, Turkey's effect is marginal, and can be understood from examination of Figure 6. The null hypothesis is that the Turkey coefficient is zero, which would imply that its effect on the model is the same as for Bahrain, for which all nation variables are zero. The line for Bahrain is decidedly offset from the centre and the line for Turkey is the furthest away in the other direction. Hence, the marginal significance of Turkey can be seen as an artifact of the nation that is selected for the null hypothesis.

We did not consider products of *ageYrs* with individual nations as predictor variables because that effectively allows different slopes for different countries. This essentially reduces to one simple linear regression per purchasing nation. The sample size as a whole is already much smaller than desirable without separating the data into smaller nation-specific samples. There would not be enough data for such a model.

Note the pathological effect of incrementally adding the highly correlated *relYrNew* by contrasting model 3 with 1 and 5 with 6. The apparent MSE increases despite the addition of a predictor. Theoretically, the optimizer should be able to find the better *relYrNew*-free model by "finding" a zero coefficient for *relYrNew*. We believe that the fact it doesn't could be related to the inaccuracy and ill-conditioning that can result from multicollinearity. Similar effects can be seen in the incremental addition of the correlated *usd2012new* by comparing model 4 with 1. Since the correlation of *usd2012new* with age is moderate, however, we decided to keep it and apply remedial techniques, to be discussed next.

4.2 Regularization

The coefficients from regression can sometimes fluctuate a lot depending on the data sample. One cause of this is multicollinearity. Other causes include a sample size that is too small, or the flip-side of that problem, too man predictors in the model for the number of data points available, thus leading to over-fitting to the specific sample. The idea behind regularization[20] is to modify the optimization function so that it doesn't just minimize the *MSE*, but it also favours small coefficients for predictors that the analyst thinks are less correlated with the dependent variable. This is done quite simply by adding the L1 or L2 norms of the corresponding coefficients to the minimization function[21], with weights provided by the analyst. Thus, coefficients are penalized according to their magnitude. If the some of the predictors in a model are products of other first order predictors, the higher order predictors often qualify for such coefficient shrinkage. If the model has only first order predictors and the variance in coefficients is due to a small sample size, then the shrinkage can also be applied to the first order coefficients. In exchange for this bias in the estimation of the coefficients, the coefficients determined are less sensitive to the specific samples. The coefficients are not as well fitted to the particular sample from which they are derived because the minimization function does not just consist of the MSE. However, the aim is for the model to perform well on new data due to the shrinkage of undue dependencies.

Three well known schemes for regularization are[22] *Ridge regression*, which uses the L2 norms; *least absolute shrinkage and selection operator (Lasso) regression*, which uses L1 norms; and *Elastic Net* regression, which uses a linear combination of L1 and L2 norms. Regularization that leans heavily

[20] References for regularization abound on the internet, but intuitive introductory material can be found in references [29, 30].

[21] The minimization function is taken to be the sum of the square of the residuals rather than the MSE.

[22] Chapter 4 of reference [31] provides a good introduction to these regularization methods. In the literature, the terms *Ridge*, *Lasso*, and *Elastic Net* appear in capitalized and uncapitalized forms, modifying both "regularization" and "regression". Elastic Net is sometimes one word, and is sometimes used as a noun.

toward Lasso regression can yield simpler regression models because of the known tendency to drive some of coefficients to zero.

We applied Elastic Net to model 6 (all predictors except for the highly correlated *relYrNew*). For comparison, we also applied Elastic Net to model 5, which consists of all predictors including *relYrNew*. The tool used was the *lasso* function in Matlab®'s statistics toolbox[23]. It penalizes all the predictor coefficients in the same way, based on a pair of hyperparameters described below. Unfortunately, *lasso* does not generate p-values for the estimated coefficients. However, it does estimate the MSE that can be expected when applying the resulting regression model to new data. For a given set of hyperparameter values, the MSE is estimated by subjecting the model generation process to cross-validation [32, 33] using the sample data. Specifically, the regression coefficients are determined from some of the data (thus *training* the model) and tested it on the rest. Data that is set aside for testing, however, is done so at the expense of training data, which is problematic for small sample sizes. Such a situation favours *leave-one-out* (LOO) cross-validation, which trains the model using $n-1$ out of n data and tests the model using only the one remaining data point. n iterations of model training and testing are performed, with each data point taking its turn as the test datum. The MSE is estimated from the resulting n prediction errors.

Matlab®'s *lasso* estimates the MSE in this way for a range of values in one of the two hyperparameters that control coefficient shrinkage, generating an MSE estimate for each case. It is simple to automate a sweep of the other hyperparameter. The best regression model is the one generated from the hyperparameter values that yield the minimum expected MSE. In *lasso*, the hyperparameters for Elastic Net are dubbed λ and α:

1. In generating a penalty, λ is the weighting factor applied to the linear combination of L1 and L2 norms for each coefficient. A larger λ means more shrinkage.
2. α linearly controls the weighted average between the L1 and L2 norms of each coefficient. $\alpha = 0$ specifies a completely L2 norm and no L1 norm; $\alpha = 1$ specifies the reverse.

Figure 7 shows the search for a good combination of hyperparameters with which to generate the regression model, i.e., $\{\lambda, \alpha\}$ values that minimize the expected MSE. For each λ value, error bars show the standard error of the MSE, as determined from the n LOO tests. The smallest MSE and corresponding λ are marked in green. Since the object of regularization is often to reduce the number of coefficients through shrinkage, *lasso* also shows in blue the largest lambda whose MSE is within one standard error of the smallest. We can see that best MSE is relatively impervious to the value chosen for α as long as there is a search over λ. In fact, the plots show that MSE bottoms out as $\lambda \to 0$, suggesting that regularization is not needed.

The results of these searches are shown in the top half of Table 10 (page 20). *DF* refers to the degrees of freedom, which in this context means the number of non-zero coefficients. All six predictors were present in all four of the $\{\alpha, \lambda\}$ combinations found, so none of the coefficients were shrunk to zero. This is not surprising considering the small values of λ and the minimal degree of regularization. *SE* refers to the standard error of the MSE estimate from LOO cross-validation. A notable feature of the Elastic Net results is that the coefficients do not change much, with the small exception of $\alpha = 0.01$,

[23]Matlab® version 2013a.

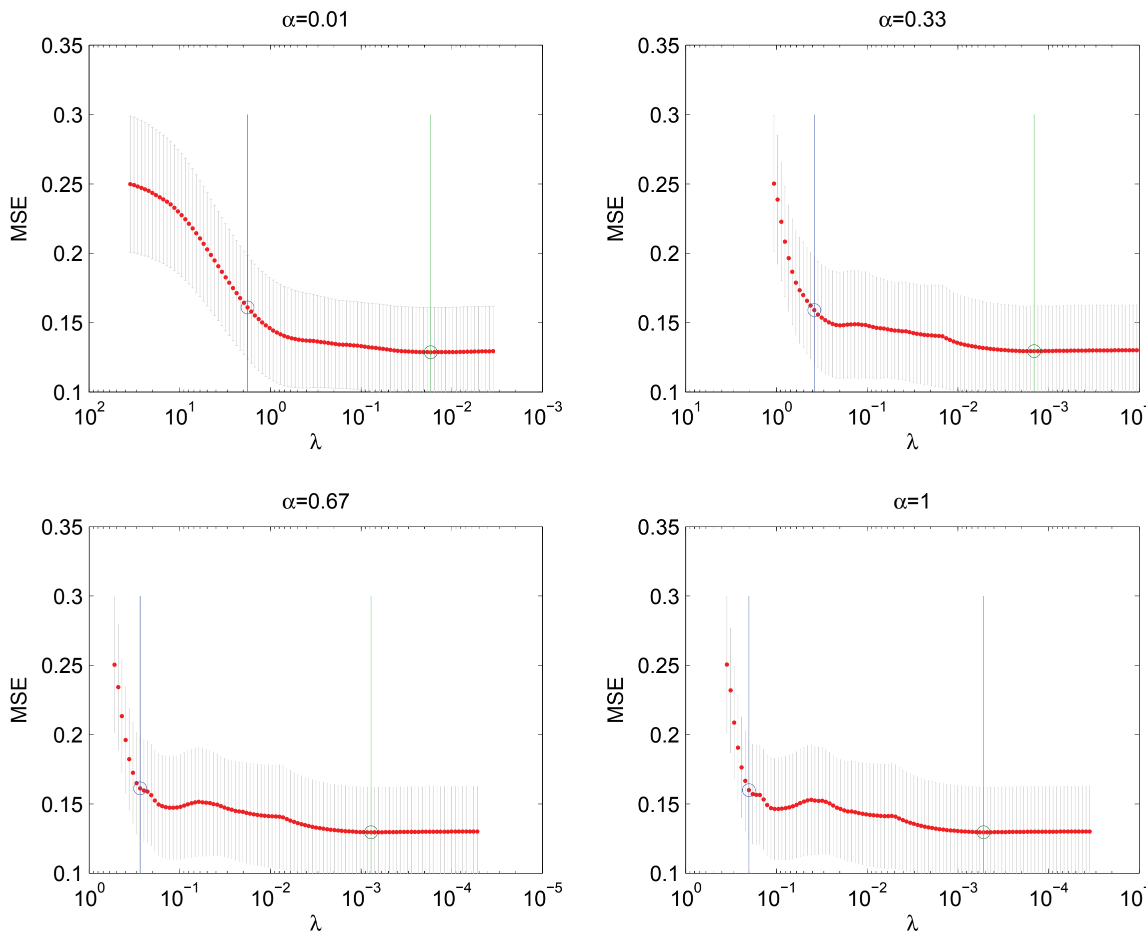

Figure 7: *MSE estimates for Elastic Net regression model with relYrNew excluded.*

i.e., when the regularization scheme leans toward heavily Ridge rather than Lasso. For $\alpha = 0.01$, the age coefficient diminishes slightly, the coefficient for Poland jumps by 20%, and the intercept drops by 6%.

The figures in the top half of Table 10 should be compared with model 6 in the preceding Table 9 on the same page, which Elastic Net was applied to. Aside from the results for $\alpha = 0.01$, the regularized coefficients are quite close to those for the unregularized model 6 in Table 9. Elastic Net's MSE estimates for new data from LOO cross-validation are about 24% higher than the root mean square of the residuals for the entire sample in model 5, but these are not exactly the same metrics. Besides, as discussed above, the regularized model is not suppose to minimize the MSE for this particular sample.

The regularization of model 6 can be contrasted with the regularization of model 5, which includes the highly correlated *relYrNew*. Figure 8 shows the hyperparameter search for regularizing model 5, the results of which are shown in the bottom half of Table 10. With the inclusion of *relYrNew*, Figure 8 shows that significant degrees of regularization do improve the MSE estimates, which demonstrates the mitigation of the effects of multicollinearity. The nontrivial values of λ in Table 10 show that significant regularization is being relied on. From the zero coefficients, we see that variables are progressively shrunk out of the model as $\alpha \to 1$ and the regularization scheme leans toward Lasso; the DF decreases accordingly. It is interesting to see that the age coefficient is smallest when the *relYrNew* coefficient is non-zero; this is the kind of trade-off that one expects from correlated predictors. Note also that *relYrNew* is one of the first predictors to be shrunken out of the model. Due to the re-apportioning of emphasis on the different coefficients as $\{\alpha, \lambda\}$ change, it should not be surprising to see that the coefficients change a lot between $\{\alpha, \lambda\}$ values. Hence it is difficult to draw a meaningful comparison with the unregularized model 5. For some coefficients, in fact, the signs of the coefficients differ between the two models. As they were for the case in which *relYrNew* was excluded, the MSE estimates for the regularization of model 5 is higher than the root mean square of the residuals in model 5, and this disparity is slightly larger when *relYrNew* is included.

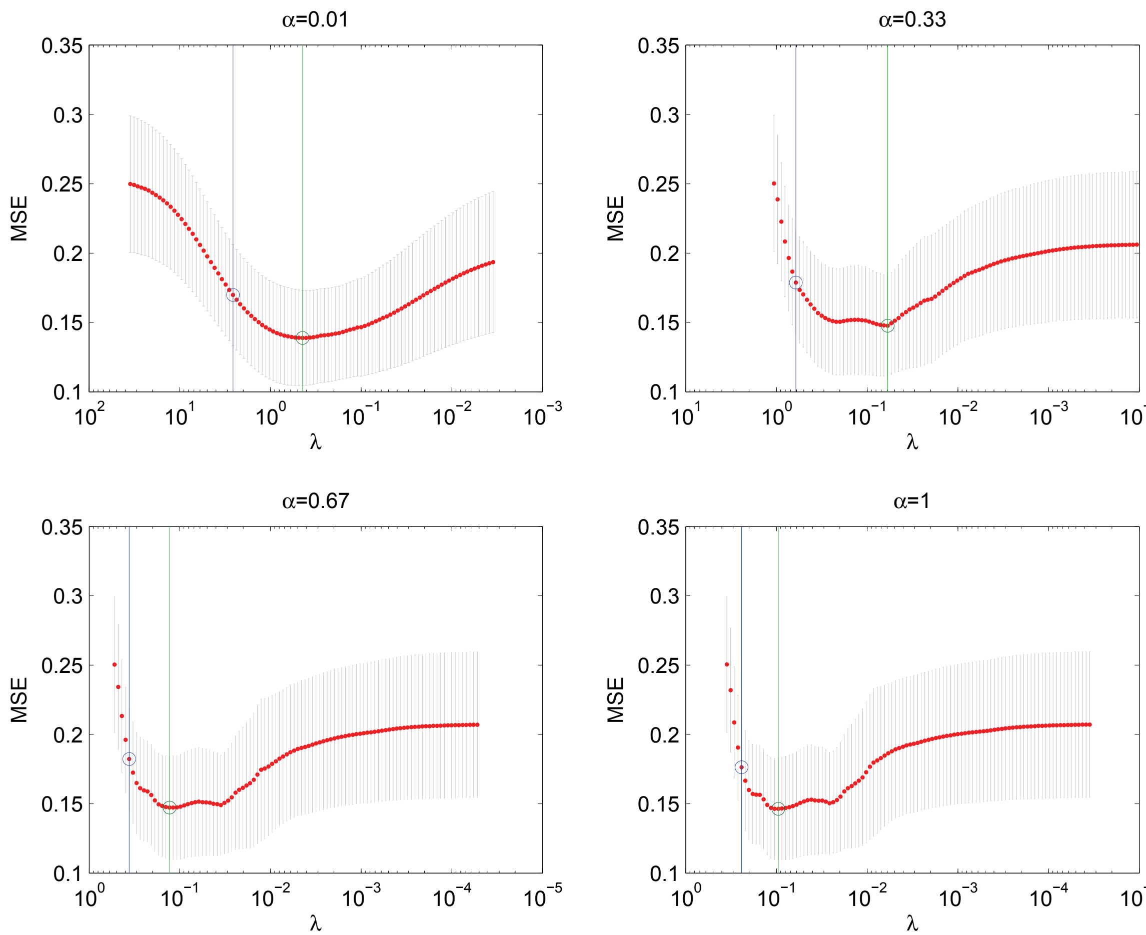

Figure 8: *MSE estimates for Elastic Net regression model with relYrNew included.*

As a closing note to this sub-section, it should be mentioned that improved regularization of models with categorical predictors has been the topic of much recent research. One area of emphasis has been to avoid treating dummy variables as completely separate predictors, since they are actually mutually exclusive categories for one categorical variable, i.e., they are *grouping variables*. From a cursory scan of the literature, Elastic Net initially seemed like it made accommodations for this by treating groups of related variables in a similar fashion. However, the kind of relationship that results in such grouping is correlation [34] rather than the mutual exclusion between the dummy variables that correspond to a categorical variable[24]. A well-known recent extension of Lasso, *group lasso* [35], was developed for models that include categorical variables. How best to apply it has been the topic of ongoing research that lies beyond the scope of this study, and as one might expect for a research frontier, many other variations have been proposed. Due to the common-ness of categorical factors in decision support, it is an area of which it would be worthwhile to keep abreast.

4.3 Impact of different inflation data

As described in Section 3.1, budget year prices were inflated to 2012 USD using inflation rates based on the CPI. Other sources were found at a later date, e.g., on an SCN TOA index. A subjective comparison of the latter with the former showed them to be similar, and it was unclear whether the sector specific SCN TOA inflation was more appropriate. As described, the latter may be more suitable for cases in which funds are inflexibility earmarked for the acquisition of naval platforms (but with enough flexibility to decide whether to buy new or used). Hence, it is prudent to check how much the calculations could vary due to the choice of the CPI inflation.

Some insight into this can be gained by comparing how the two sets of inflation compound over the period of interest, as shown in the left half of Figure 9. The curves reflect how the budget year price of a hypothetical \$1 item changes through the years due to the two sets of inflation rates. This must be scaled by the actual price of a specific item of interest (hundreds of thousands of dollars for an FFG-7), but that is just a multiplicative factor. The factor by which the price appreciates is still represented by the curves in Figure 9. The disparity between the curves provides a rough idea of the scale of the price divergence that is possible under the two inflation data sets. For the purpose of calculating depreciation, the actual divergence depends on when an item was new (t_{New}) and when it was sold used (t_{Used}), which defines the period of compounded inflation that affects the ratio $Cost_{\text{Used}}/Cost_{\text{New}}$. For convenience, one may convert $Cost_{\text{Used}}$ and $Cost_{\text{New}}$ to budget year dollars for a year outside of $[t_{\text{New}}, t_{\text{Used}}]$, as was done in this study, but the multiplicative effect of inflation from the years outside of $[t_{\text{New}}, t_{\text{Used}}]$ will be the same for both $Cost_{\text{Used}}$ and $Cost_{\text{New}}$. Thus, inflation outside of $[t_{\text{New}}, t_{\text{Used}}]$ does not affect $Cost_{\text{Used}}/Cost_{\text{New}}$.

Figure 9 (right hand side) also plots the offset of the two compounded inflation curves from their mean (not shown), normalized by this very same mean. In effect, since we don't have concrete and specific information with which to choose one curve over the other, we choose the average between the two curves as a reference curve with which to normalize differences. The offset of each curve from the average, which is half the difference between the two curves, is then expressed as a percentage of the average. Figure 9 shows that this normalized offset peaks at 8%. Since the average age at which FFG-7s are sold used is 18.5 years, this 8% divergence represents a typical annual multiplicative factor of roughly $1.08^{1/18.5}$, or approximately 0.4%/year. Such an offset is within the 68% confidence interval for the depreciation (Table 6, page 15).

[24] The study of the correlation between dummy variables is beyond the scope of this effort.

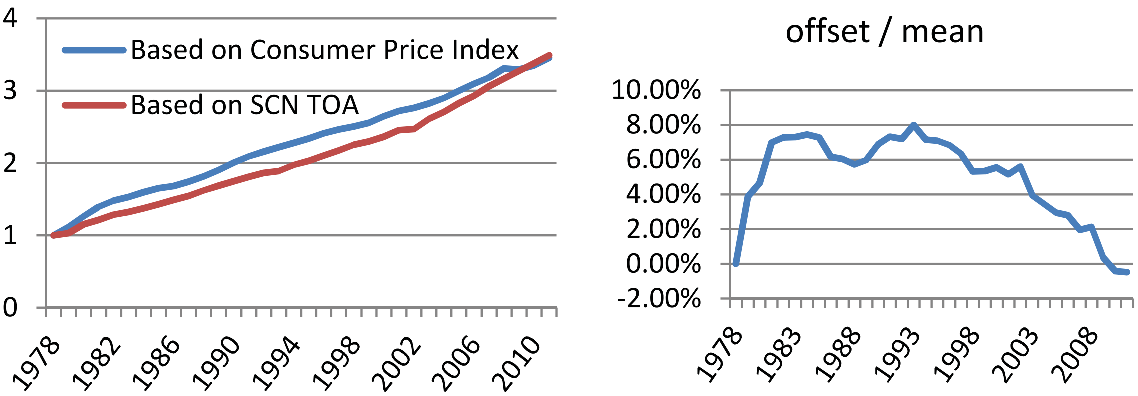

Figure 9: *Comparison of compounded inflation from two data sets.*

More explicitly, let d_1 be the naive annual depreciation calculated based on a single ship that is sold used after 18.5 years, i.e. $1 - d_1 = (Cost_{\text{Used}}/Cost_{\text{New}})^{1/18.5}$, where $Cost_{\text{Used}}$ and $Cost_{\text{New}}$ are the apparent real costs in terms of whatever year and currency is adopted as a reference. Assume that this was determined using the CPI inflation figures, and imagine that we later found out that the SCN TOA inflations were deemed to be more suitable, for whatever reason. According to Figure 9, the figure $Cost_{\text{Used}}$ is actually 8% lower (worst case scenario) than initially thought when calculating d_1. If we designate d_2 as the correct depreciation, then it should satisfy: $1 - d_2 = (\xi Cost_{\text{Used}}/Cost_{\text{New}})^{1/18.5}$, where $\xi = 0.92$ to account for the 8% drop. Hence,

$$\frac{1 - d_2}{1 - d_1} = \xi^{1/18.5} = 0.9955 \, . \tag{6}$$

Using $d_1 = 8.4\% = 8.4 \times 10^{-2}$ from Table 6 (page 15), we get $d_2 = 8.8 \times 10^{-2}$, which is about 0.4 percentage points greater than d_1 [25]. Basically, the SCN TOA inflation accounts for less of the budget year drop in price, so depreciation must be responsible for slightly more of it. The bottom line, however, is that the worst case divergence of budget year price has a minor effect on the real depreciation when compared to the uncertainty due to sampling noise. Because of compounding, this remains true even if we use the full difference between compounded inflation curves to calculate offset/mean in Figure 9 instead of the half difference (the worst case offset becomes 16% instead of 8%).

4.4 Sensitivity analysis summation

From multiple regression on various various sets of predictor variables, the age of a ship is the only statistically significant predictor of depreciation in the unregularized models. Predictors that were obviously correlated were vetted away. Treating the purchasing nation as a categorical variable yielded predictions that graphically seemed to fit better with this specific data set, but the p-values

[25] In fact, (6) implies that $d_2 - d_1 = (1 - 0.9955)(1 - d_1) = 0.45\%(1 - d_1)$, which is slightly less than 0.45% for plausible depreciations d_1 and d_2 regardless of the exact values.

indicate that the nation specific dependencies are not significant. For the full model (all predictors excluding the predictor corresponding to the highly correlated year-when-new), regularization did not change the model much and was not needed.

The full model included a predictor for the year when a ship was new (relative to 15 December 1979), which was found to have significant negative correlation with the age of the ship when it was sold second hand, i.e., ships built later were sold at an earlier age. This introduced significant multicollinearity into the regression, which yielded anomalous results. The use of nontrivial degrees of regularization improved the model in terms of the residual errors that can be expected when the model is used on new data, though it was still higher than for the regularized models in the absence of significant multicollinearity. Regularization also shrunk various predictors out of the model, as expected, depending on the hyperparameter values used. The re-apportioning of dependence among the predictors yielded coefficients that varied greatly between models that were generated from different hyperparameter values, with some coefficients having the opposite sign from the un-regularized model.

Due to the lack of p-values from the generation of regularized models, it is not clear whether the coefficients in the regularized models are any more significant that those in the unregularized models. We expect that they are similar for the full unregularized model, minus the predictor for the relative year-when-new, since the regularized and unregularized models are so similar. Hence: (i) none of the unregularized multiple regression models are warranted over simple linear regression; (ii) the regularized models in the absence of significant multicollinearity are probably not much different, at least based on the limited amount of testing thus far[26]; and (iii) there is no compelling reason to recommend the regularized model with significant multicollinearity. Concerning the final point, we are not aware of any information suggesting that significance of predictors is improved by regularizing a model with multicollinearity, either intentionally introduced or not.

Regarding sensitivity of the depreciation modelling to the inflation data, the difference between inflation based on the CPI versus SCN TOA index were too small to have an appreciable effect.

[26]Limited by the small data set

5 Discussion

5.1 Depreciation as per used sales differs from characteristic degradation of platform

As mentioned, the parts hulk was excluded in the normality tests. Whether a parts hulk should be included in a model depends on whether we wish to model the depreciation of a platform in general or to model only the instances of a platform that are resold for operational use. For this study, and likely in all future studies, the CAF is interested in second hand platforms for operational use rather than for parts. Even in that case, however, there may be good reasons for getting a good idea of how the platform of interest depreciates in general. One reason is the possibility that the buyer may end up buying a platform that becomes a parts hulk not by design. As an example of a used purchase that did not go as planned, consider Canada's purchase of four Victoria class submarines from the U.K. in 1998 [36]. One of the submarines, the HMCS Chicoutimi, was beset by a fatal fire en route to Canada and left adrift. Restoration of the Chicoutimi was completed 10 years later and it was recently being returned to the navy for sea trials [37]. Though Chicoutimi avoided the fate of a parts hulk, its return to service was anticipated to be at a reduced level[27], and it is not hard to see how unanticipated setbacks could relegate a used purchase to the role of a parts hulk.

There is a second, not entirely unrelated, perspective with which one can view the case for including parts hulks in analyses, possibly as a point of comparison for analyses without part hulks. Used sales do not include vessels that are in too poor condition to sell, nor vessels that the selling nation keeps for its own use, which could be those that age best. Therefore, one should be wary of the possibility that, unlike for a commodity like a car, examining used sales for major platforms could yield a very limited indication of how a platform truly degrades with age. In other words, the characteristic degradation of a platform differs from trends in used sale prices. If the true degradation characteristics can be determined, such information would provide a sense of risk in buying used, is helpful for identifying and prioritizing options and establishing initial positions in price negotiations, and can inform decision makers on the level of diligence needed in assessing the condition of the used platforms. While characteristic platform degradation cannot be determined from used sales data alone (such data exists only for used sales, which is the very problem), it is possible to compound the information shortfall by dismissing the part hulk data points outright. We chose to address this by comparing analyses with and without the parts hulk, but taking the latter as the default case.

5.2 Physical-age and design-age contributions to depreciation

In order to get insight into the depreciation behaviour of a particular platform, it is necessary to delve into the contributing factors of depreciation rather than simply treating depreciation as a black box process. Without the parts hulk, the annual depreciation of roughly 8.4%/year that is indicated by the used sale prices includes contributions from loss of utility relative to "state of the art" in addition to the difference due to "old versus new". We will refer to these components as design-age depreciation and physical-age depreciation, respectively. The former is due to the fact that, even if platforms are built new, designs from many years past are less able to meet current military requirements due to factors such as increased expectations due to advancing technology, the evolving capabilities of

[27] "Chicoutimi will...be restricted to shallow-water diving for the foreseeable future" [37].

adversaries, the evolving nature of the responsibilities that militaries are called upon to fulfil, and new concepts of operation that require more technology-intensive coordination and communication. For example, consider a ship that is ready for its first ever commissioning. If it could be hypothetically kept out of operational service and prevented from physically degrading (rusting, wear and tear, etc.), design-age depreciation captures the fact that the technology and design would still become obsolete and misaligned with evolving demands, causing its value to decline despite the fact that it does not physically age.

Design-age depreciation is clearly implied in Pugh's source book of defence equipment costs [38], which shows that the real cost of the individual classes of new platforms "built to the latest design standards and now entering service" typically increase from 1% to 5% (and sometimes up to 10% and more) per year. The physical-age and design-age components of depreciation are conflated in our data and are difficult to disentangle. Hypothetically, if a nation was still building new FFG-7s with capabilities not differing much from the existing FFG-7s *at the time when they were new*, we would have data points for design-age depreciation without physical-age depreciation – potentially, these effects could then be separated. Compounding the conflation of the two contributors to depreciation is the fact that, in response to evolving military requirements, militaries can typically be expected to combat design-age depreciation through betterments throughout the life cycle of platforms. Hence, we can expect only part of the design-age contribution to be captured in typical depreciation data.

5.3 Attempted extraction of design-age depreciation

We attempted to extract a rough estimate of design-age depreciation from the FFG-7 used sales data set. As an approximate value range, the annual growth in the real cost of a new vessel is attributed to technology and design improvements. We emphasize that the purpose of considering the increasing cost of newly built vessels is not to expand the scope of this study to include the option analysis of buying new; the purpose is to estimate the design-age contribution to depreciation. The percentages for annual growth in real cost for new vessels are taken to be year-to-year reductions in real cost if year-to-year improvements in technology and design are not implemented. Thus they serve as a proxies for design-age depreciation. This counter-factual analysis[28] merely augments the main depreciation analysis for insight, so we did not undertake a second wave of data gathering to obtain costs-when-new for FFG-7s that were not sold used[29].

The costs of new FFG-7s in 2012 USD were plotted against the initial service date (Annex E). Following Pugh's model of annualized growth in real cost, we log transformed the costs in order to use linear regression to determine the exponential growth with time. This mirrors what was done to calculate FFG-7 depreciation, except that the aim here is to characterize growth rather than decay. Due to the short time interval spanned by the initial service dates (1979~1982), the estimated time dependence from regression is very rough. In Figure E.1, the slope of -0.0371 yields an estimated annual real cost "growth" for new FFG-7s of $(e^{-0.0371} - 1)100\% = -3.6\%$[30]. It was evident that if Pugh's real cost growths applied to FFG-7s, then our sample was too small and our time window too

[28] A counter-factual study examines scenarios other than those that actually occurred.

[29] FFG-7s in active service, disposed of, or simply decommissioned.

[30] For such a rough check as this, we did not attempt to account for fact that the resulting net growth includes the effect of decreasing cost with rank in class.

narrow for its detection. For our data set, therefore, design-age depreciation could not be extracted using real cost growth as a proxy.

5.4 Accounting for sampling noise in design-age depreciation

The general trend of real cost growth was not detected in our sample, but it would be prudent to determine if the discrepancy can be attributed to sampling noise. We did this by comparing our "noisy" estimate of "growth" rate to the most appropriate categories of ships in Pugh's study (Pugh refers to ship categories as *classes*). To identify the applicable classes, note that the ostensible function of the FFG-7 is air defence and antisubmarine warfare (ASW) [2], and its full load displacement of 4200 tons is approximately midway between these two "Pugh" warship classes (2500 and 7000 tons). Hence, we associate the FFG-7 with both these classes. Fortunately, the normalized new costs for these two classes are quite close to each other compared to the other vessel classes, and Pugh's annual growth in real cost is 2% for both classes. Thus, we take the null hypothesis to be 2%/year growth and determine the 95% confidence interval using the standard error of the estimated slope from the above regression. This yields a 95% confidence interval of [-6.7%,11.5%]/year growth in real cost[31], which encompasses the -3.6% estimated from the data above. The p-value for -3.6%/year "growth" is 0.195, which is high and indicates a dearth of evidence against the hypothesis of 2%/year. In short, given the sampling noise, the -3.6% real "growth" in buying new is not inconsistent with Pugh's 2%/year growth in the cost of new ships in a similar class.

In an attempt to address the broad sampling noise and the short time span in the above regression for new cost growth, we incorporated into the data sample two FFG-7s built in Australia[32] at the much later date of 1992 and a budget-year cost of US$385M each ($636M in 2012 USD). However, these two data points turned out to be quite anomalous compared to the rest of the sample, which led to less clarity on the real cost growth of buying new. A regression on the expanded sample yielded an estimated real cost growth of 5.8%/year, but the 95% confidence interval flanking the null hypothesis of 2%/year shrank to [-0.5%,4.6%]/year. Since this excludes the estimated 5.8%/year, it implies that either Pugh's quantitative generalization does not apply to FFG-7s[33] or there is something very distinguishing about these Australian FFG-7s. As a check on the data, the USD currency for this data and the fact that the cost figures were per-ship were corroborated via multiple information articles. If the data is not accurate, then a common error propagated to the multiple information sources. To get a perspective on the extremeness of this anomaly from Pugh's published trends, Annex E contains a projection of the price range of a notional new FFG-7 in 1992, outfitted to applicable "modern" standards, i.e., modern for 1992. The resulting quartiles define the 50% confidence interval for the

[31] Since the linear regression was done in the log domain, the null hypothesis is that the slope is $\ln(1 + 2\%/100\%)$/year. Conversely, upper and lower bounds on the 95% confidence interval for the slope are converted to a growth rate interval via $(e^{Slope} - 1)100\%$. A t-distribution is used to determine the upper/lower bounds.

[32] The Melbourne and the Newcastle.

[33] The data points on which the Pugh's trends are based approximately span the years [1976,2008]. This encompasses the period [1979,1992] that is spanned by the initial service dates of the FFG-7s in our sample, including the two anomalous Australian vessels in 1992. Based on the available information, therefore, the discrepancy cannot be dismissed as due to disparate time periods.

price as [$151M,$245M] 2012 USD for the notionally new FFG-7s in 1992[34]. The apparent cost of $636M 2012 USD for each of the two Australian FFG-7s in 1992 lies well outside this range.

[34] The quartile-based interval provides a quick and rough check of the Australian data points against the projected cost range. Other confidence intervals were not used because the quartiles are highly asymmetric about the median both with and without a log transformation, thus ruling out a normal distribution for the null hypothesis.

6 Conclusions and recommendations

The depreciation of used FFG-7 frigates was analyzed as part of a plan to study the life cycle trade-offs in cost of buying major platforms second hand. While the study was put on hold in part due to lack of data, some initial planning of the database has been documented here, and specific approaches and challenges in modelling a change of ownership have been identified in case such a study is pursued again. New and used cost data for used platform sales is expected to be challenging to obtain.

Data collection was performed for both FFG-7s and Kortenaers from entirely open source, with assistance from DHH, OSINT, and U.S. DSCA. The process was iterative, involving repeated deconflictions between the developing database and newly encountered data. Judgement was exercised in determining new and used costs and transition dates not only because of variety of information sources, but also due to the various phases that might be involved in decommissioning operations, transferring a ship to another nation, and making them operational again. Some uncertainty in the data arises from the fact that frigates sold used were often accompanied by value enhancements. For traceability, as much of the source information as was practical has been captured in a repository.

While the database was fairly well populated in general, there was a lack of new and used cost data with which to analyze the Kortenaers. Seventeen of the FFG-7s had sufficient cost data, which makes a small sample. All costs were inflated to 1 January 2012 using inflation rates based on the U.S. CPI, which were found to be in line with shipbuilding inflation figures. Depreciations were found to roughly fit an exponential decay model in the sense that the residuals were normally distributed in the log domain. After an initial drop in value of 21%, the depreciation of 8.4% / year was found to explain 56% the variance, making it a very rough determinant of used ship cost. This explanatory power dropped to 41% with the inclusion of an outlier associated with one frigate that was purchased as a parts hulk. These levels of explanatory power resulted from the use of just one input variable in the model.

The validity of the regression performed on used sale costs relies in part on normally distributed residuals. The normal probability plot showed a good fit, devoid of indications of skewness and kurtosis. Both the Anderson-Darling test and Doornik-Hansen omnibus test were meant for small sample sizes, and both yielded no appreciable evidence against normality. The Doornik-Hansen test did not reveal any significant deviations from normality in the form of skewness or excess kurtosis. Some of the theory and intuition behind normality testing explored in annexes, including a method to determine confidence bands on a normal probability plot for small sample sizes.

Significant sensitivity testing was done by exploring multiple regression models with and without regularization, and determining the representative impact of an alternate set of inflation data on the model. Multiple regression was not found to be warranted as the other variables did not show any significance, and the simple linear depreciation model was impervious to the choice of inflation data.

The estimated annual depreciation includes physical-age and design-age contributions. An estimate of the design-age component was attempted, but it was not detectable with the small sample size and narrow time window. However, no evidence was found from our sample data against a published comprehensively derived general trend of a 2%/year growth in real cost for the applicable vessel classes bought new. A higher fidelity estimate was attempted by including two modern Australian FFG-7s in the data set, but they were found to be extreme outliers for reasons unknown. It is possible for more a conclusive trend to be determined from more price data for newly-bought FFG-7s.

The statistical estimation of depreciation is a valuable input into life cycle costing in comparing the options of buying used versus new. This approach can be applied to other major platforms, such as support ships, destroyers, other frigates, vehicles, and aircraft. Furthermore, with appropriate study into the modelling of the disruptive effects in changing owners, there is the potential to adapt stochastic O&M cost modelling and data mining for purchase cost estimation to second hand purchases. These types of financial analysis are an essential part of a comprehensive comparison that ideally encompasses additional factors such as operational suitability, proficiency, and timely operationalization.

The applicability of depreciation modelling is not limited to the buying of second hand platforms by the CAF. It can also help determine an initial price range if the CAF was to consider selling assets, e.g., the Halifax-class frigates.

References

[1] Carter, B., Directorate, Materiel Group Operational Research (2012), Major Equipment Procurement Study: New Versus Used. Unpublished outline of study plan. Completion of the report on this study is slated for 2014.

[2] Oliver Hazard Perry-class frigate (online), http://en.wikipedia.org/wiki/Oliver_Hazard_Perry-class_frigate (Access Date: 10 July 2013).

[3] The Auditor General of Canada (2009), 2009 Fall Report of the Auditor General of Canada to the House of Commons, Ch. 5: Acquiring Military Vehicles for Use in Afghanistan, Office of the Auditor General of Canada. http://www.oag-bvg.gc.ca/internet/English/parl_oag_200911_05_e_33206.html (Access Date: 16 October 2013).

[4] Gubler, A. (May 2008), Defence Business: Sirius supply stalwarts Success (online), Australian Defence Magazine, http://www.australiandefence.com.au/2BB89700-F807-11DD-8DFE0050568C22C9 (Access Date: 16 October 2013).

[5] Kaluzny, B. L., Barbici, S., Berg, G., Chiomento, R., Derpanis, D., Jonsson, U., Shaw, R. H. A. D., Smit, M. C., and Ramaroson, F. (2011), An Application of Data Mining Algorithms for Shipbuilding Cost Estimation, *Journal of Cost Analysis and Parametrics*, 4(1), 2–30.

[6] Kaluzny, B. L. and Shaw, R. H. A. D. (2011), A Parametric Approach to Cost Estimation during Options Analysis: An Independent Cost Estimate of a Joint Support Ship, (DRDC CORA TM 2011–068) Defence R&D Canada – CORA.

[7] Desmier, P. E. (2010), The Optimal Time to Upgrade or Replace the Light Armoured Vehicle (LAV III) Fleet, (DRDC CORA TM 2010-101) Defence R&D Canada – CORA.

[8] Maybury, D. W. (2000), Optimal Fleet Replacement and Forecasting Under Uncertainty: The CP-140A Arcturus Martime Surveillance Aircraft, (DRDC CORA TR 2009-001) Defence R&D Canada – CORA.

[9] Kortenaer-class frigate (online), http://en.wikipedia.org/wiki/Kortenaer-class_frigate (Access Date: 26 June 2013).

[10] Kaluzny, B. L. (2011), The Unit Recurring Flyaway Cost of a Canadian Joint Strike Fighter, (DRDC CORA TM 2011–200) Defence R&D Canada – CORA.

[11] Arena, M. V., Younossi, O., Brancato, K., Blickstein, I., and Grammich, C. A. (2008), Why Has the Cost of Fixed-Wing Aircraft Risen? A Macroscopic Examination of the Trends in U.S. Military Aircraft Costs over the Past Several Decades, (Technical Report MG696) RAND Corporation, Anata Monica, CA.

[12] Moral hazard (online), http://en.wikipedia.org/wiki/Moral_hazard (Access Date: 4 October 2013).

[13] Dixit, A. K. and Pindyck, R. S. (1994), Investment Under Uncertainty, Princeton, NJ: Princeton University Press.

[14] Martingale (probability theory) (online), http://en.wikipedia.org/wiki/Martingale_(probability_theory) (Access Date: 10 October 2013).

[15] McMahon, T. (April 2013), Historical Inflation Rate [data from 1913 to the present] (online), http://inflationdata.com/inflation/Inflation_Rate/HistoricalInflation.aspx (Access Date: 9 July 2013).

[16] Bohdan L. Kaluzny, author of [5]. Personal communication, 8 July 2013.

[17] Pfeiffer, L. (2010), Historical SCN Total Obligational Authority Index. Historical/Projected Rates based on Shipbuilding Inflation (spreadsheet). Prepared as a member of NAVSEA 05C – Cost Engineering and Industrial Analysis Division.

[18] Naval Center for Cost Analysis (2013), Joint Inflation Calculator (online), http://www.ncca.navy.mil/tools/inflation.cfm (Access Date: 17 July 2013). The calculator is updated with new versions annually.

[19] Bureau of Labor Statistics, Bureau of Labor Statistical Indexes: Material Index for Steel Vessel Contracts (online), NAVSEA 05C – Cost Engineering and Industrial Analysis Division, http://www.navsea.navy.mil/Organization/SEA05C/indices.aspx (Access Date: 9 July 2013).

[20] BLS's Producer Price Index (PPI) for newly constructed military self-propelled ships (online), http://www.bls.gov/ppi/data.htm (Access Date: 28 June 2013). The site is complex and it may not be clear how to access the above data. The data was accessed via the "Commodity Data" database, multi-screen data search, selecting both seasonally adjusted and not, choosing Group 14 (Transportation equipment), then Group 3102[01], and retrieving the resulting then Group 3102[01], and retrieving the resulting WPU14310201 file. The full commodity code 14010201 was found at http://www.bls.gov/ppi/#tables under "Commodities, all levels" (ftp://ftp.bls.gov/pub/special.requests/ppi/comrlp13.txt).

[21] Mann, P. S. (1998), Introductory Statistics, 3rd ed, New York: John Wiley & Sons, Inc.

[22] Totton, N. and White, P. (July 2011), The Ubiquitous Mythical Normal Distribution (online), Bristol, UK, http://www.nbt.nhs.uk/sites/default/files/filedepot/incoming/NormalDistribution.pdf (Access Date: 10 October 2013). Organizational affiliations: University of the West of England, North Bristol NHS Trust.

[23] Tukey, J. W. (1977), Exploratory Data Analysis, Reading, MA: Addison-Wesley.

[24] Determine if data are normally distributed using SPSS: Assessing Data Normality (Gaussian, Bell Shaped Curve) (online), http://www.stattutorials.com/SPSS/TUTORIAL-SPSS-Assess-Normality.htm (Access Date: 10 October 2013). Copyright 1996-2008. Website: SPSS Tutorials for Statistical Data Analysis.

[25] Henry C. Thode, Jr. (2002), Testing for Normality, New York, NY: Marcel Dekker, Inc.

[26] Doornik, J. A. and Hansen, H. (2008), An Omnibus Test for Univariate and Multivariate Normality, *Oxford Bulletin of Economics and Statistics*, 70, 927–939.

[27] Multicollinearity (online), http://en.wikipedia.org/wiki/Multicollinearity (Access Date: 7 March 2014).

[28] How To Find Relationship Between Variables, Multiple Regression (online), Tulsa, OK: StatSoft, Inc., http://www.statsoft.com/Textbook/Multiple-Regression (Access Date: 7 March 2014).

[29] Ng, A., Regularization (online), http://www.holehouse.org/mlclass/07_Regularization.html (Access Date: 7 March 2014). Course notes.

[30] Ng, A., Exercise 5: Regularization (online), http://openclassroom.stanford.edu/MainFolder/DocumentPage.php?course=MachineLearning&doc=exercises/ex5/ex5.html (Access Date: 7 March 2014). Exercise notes.

[31] der Kooij, A. J. V. (2007), Prediction Accuracy and Stability of Regression with Optimal Scaling Transformations, Ph.D. thesis, Leiden University, Leiden, NL.

[32] Mohamad, M. S., Classification (online), Kuala Lumpur, MY: Universiti Teknologi Malaysia, http://web.utm.my/aibig/images/saberi/genome_mining/classification.ppt (Access Date: 7 March 2014). Notes for Genome Mining course SCB 4233, http://web.utm.my/aibig/people/mohd-saberi-mohamad/teaching/genome-mining.html.

[33] Refaeilzadeh, P., Tang, L., and Liu, H., Cross-Validation (online), http://leitang.net/papers/ency-cross-validation.pdf (Access Date: 7 March 2014).

[34] Zou, H. and Hastie, T. (2005), Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society, Series B*, 67, 301–320.

[35] Yuan, M., Lin, Y., and Lin, Y. (2006), Model selection and estimation in regression with grouped variables, *Journal of the Royal Statistical Society, Series B*, 68, 49–67.

[36] Upholder/Victoria-class submarine (online), http://en.wikipedia.org/wiki/Upholder/Victoria-class_submarine (Access Date: 11 October 2013).

[37] (January 2014), Rebuilt HMCS Chicoutimi submarine to return to navy (online), The Canadian Press, http://www.cbc.ca/news/canada/nova-scotia/rebuilt-hmcs-chicoutimi-submarine-to-return-to-navy-1.2486251 (Access Date: 10 March 2014).

[38] Pugh, D. G. (2009), Source book of defence equpipment costs, Clapham, Bedfordshire: D. G. Pugh.

[39] Simple linear regression: Normality assumption (online), http://en.wikipedia.org/wiki/Simple_linear_regression#Normality_assumption (Access Date: 18 July 2013).

[40] Normal probability plot (online), http://en.wikipedia.org/wiki/Normal_probability_plot (Access Date: 11 July 2013).

[41] Q-Q plot (online), http://en.wikipedia.org/wiki/Q-Q_plot (Access Date: 11 July 2013).

[42] Inverse transform sampling (online), http://en.wikipedia.org/wiki/Inverse_transform_sampling#The_method (Access Date: 11 July 2013).

[43] Law, A. M. (2007), Simulation Modeling & Analysis, 4th ed, Industrial Engineering and Management Science, Boston: McGraw-Hill.

[44] Brandt, S. (1999), Data Analysis: Statistical and Computational Methods for Scientists and Engineers, 3rd ed, New York: Springer.

[45] Probability integral transform (online), http://en.wikipedia.org/wiki/Probability_integral_transform (Access Date: 19 July 2013).

[46] Order statistic (online), http://en.wikipedia.org/wiki/Order_statistic (Access Date: 11 July 2013).

[47] Berthouex, P. M. and Brown, L. C. (2002), Statistics for Environmental Engineers, 2nd ed, Boca Raton, FL: CRC Press.

[48] Hahn, G. J. and Shapiro, S. S. (1967), Statistical Models in Engineering, Wiley Series on Systems Engineering and Analysis Series, New York, Chichester, Brisbane: John Wiley & Sons.

[49] Anderson-Darling test (online), http://en.wikipedia.org/wiki/Anderson-Darling_test (Access Date: 11 July 2013).

[50] Cramér-von Mises criterion (online), http://en.wikipedia.org/wiki/Cramér-von_Mises_criterion (Access Date: 11 July 2013).

[51] D'Agostino, R. B. and Stephens, M. A., (Eds.) (1986), Goodness-of-fit techniques, Vol. 68 of *STATISTICS: Textbooks and Monographs*, New York: Marcel Dekker, Inc.

[52] Stephens, M. A. (1974), EDF Statistics for Goodness of Fit and Some Comparisons, *Journal of the American Statistical Association*.

[53] Stephens, M. A. (1970), Use of the Kolmogorov-Smirnov, Cramer-Von Mises and Related Statistics Without Extensive Tables, *Journal of the Royal Statistical Society. Series B (Methodological)*, 32(1), 115–122.

[54] Kolmogorov-Smirnov test (online), http://en.wikipedia.org/wiki/Kolmogorov-Smirnov_test (Access Date: 19 July 2013).

[55] Ihaka, R., Theoretical Quantile-Quantile Plots (online), Auckland, NZ, http://www.stat.auckland.ac.nz/~ihaka/787/lectures-quantiles2-handouts.pdf (Access Date: 19 July 2013). Handouts for lecture Quantiles2 in University of Auckland course Statistics 787: Topic in Computational Data Analysis and Graphics.

[56] Gibbons, J. D. and Chakraborti, S. (2003), Nonparametric Statistical Inference, 4th ed, New York: Marcel Dekker, Inc.

[57] Razali, N. M. and Wah, Y. B. (2011), Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests, *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.

[58] Lilliefors, H. W. (1967), On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown, *Journal of the American Statistical Association*, 62(318), 399–402.

[59] Croarkin, C. and Tobias, P., (Eds.) (2012 significant update), NIST/SEMATECH e-Handbook of Statistical Methods, Ch. 1.3.5.16: Kolmogorov-Smirnov Goodness-of-Fit Test, (US) National Institute of Standards and Technology. Accessed at http://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm on 20 July 2013. Less specific, more robust URL: http://www.itl.nist.gov/div898/handbook.

[60] Meeker, W. Q. and Escobar, L. A. (1998), Statistical Methods for Reliability Data, Wiley series in probability and statistics, Ch. 6.4: Probability Plotting Positions, New York: John Wiley & Sons, Inc.

[61] Pearson, E. S. and Hartley, H. O., (Eds.) (1972), Biometrika Tables for Statisticians, Vol. II, London: Cambridge University Press.

[62] Conover, W. J. (1999), Practical Nonparametric Statistics, 3rd ed, New York: John Wiley & Sons, Inc.

[63] Makhnin, O. V., Kolmogorov-Smirnov Test (online), New Mexico Institute of Mining and Technology, http://www.nmt.edu/~olegm/484/Section2-7.pdf (Access Date: 7 September 2013). Course notes for Math 484 (Reliability and quality control).

[64] OANDA Historical Exchange Rates (online), OANDA, http://www.oanda.com/currency/historical-rates (Access Date: 31 July 2013).

This page intentionally left blank.

Annex A: U.S. inflation based on CPI

The following inflation rates based the U.S. Consumer Price Index were obtained from [15].

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
1979	9.28%	9.86%	10.09%	10.49%	10.85%	10.89%	11.26%	11.82%	12.18%	12.07%	12.61%	13.29%	11.22%
1980	13.91%	14.18%	14.76%	14.73%	14.41%	14.38%	13.13%	12.87%	12.60%	12.77%	12.65%	12.52%	13.58%
1981	11.83%	11.41%	10.49%	10.00%	9.78%	9.55%	10.76%	10.80%	10.95%	10.14%	9.59%	8.92%	10.35%
1982	8.39%	7.62%	6.78%	6.51%	6.68%	7.06%	6.44%	5.85%	5.04%	5.14%	4.59%	3.83%	6.16%
1983	3.71%	3.49%	3.60%	3.90%	3.55%	2.58%	2.46%	2.56%	2.86%	2.85%	3.27%	3.79%	3.22%
1984	4.19%	4.60%	4.80%	4.56%	4.23%	4.22%	4.20%	4.29%	4.27%	4.26%	4.05%	3.95%	4.30%
1985	3.53%	3.52%	3.70%	3.69%	3.77%	3.76%	3.55%	3.35%	3.14%	3.23%	3.51%	3.80%	3.55%
1986	3.89%	3.11%	2.26%	1.59%	1.49%	1.77%	1.58%	1.57%	1.75%	1.47%	1.28%	1.10%	1.91%
1987	1.46%	2.10%	3.03%	3.78%	3.86%	3.65%	3.93%	4.28%	4.36%	4.53%	4.53%	4.43%	3.66%
1988	4.05%	3.94%	3.93%	3.90%	3.89%	3.96%	4.13%	4.02%	4.17%	4.25%	4.25%	4.42%	4.08%
1989	4.67%	4.83%	4.98%	5.12%	5.36%	5.17%	4.98%	4.71%	4.34%	4.49%	4.66%	4.65%	4.83%
1990	5.20%	5.26%	5.23%	4.71%	4.36%	4.67%	4.82%	5.62%	6.16%	6.29%	6.27%	6.11%	5.39%
1991	5.65%	5.31%	4.90%	4.89%	4.95%	4.70%	4.45%	3.80%	3.39%	2.92%	2.99%	3.06%	4.25%
1992	2.60%	2.82%	3.19%	3.18%	3.02%	3.09%	3.16%	3.15%	2.99%	3.20%	3.05%	2.90%	3.03%
1993	3.26%	3.25%	3.09%	3.23%	3.22%	3.00%	2.78%	2.77%	2.69%	2.75%	2.68%	2.75%	2.96%
1994	2.52%	2.52%	2.51%	2.36%	2.29%	2.49%	2.77%	2.90%	2.96%	2.61%	2.67%	2.67%	2.61%
1995	2.80%	2.86%	2.85%	3.05%	3.19%	3.04%	2.76%	2.62%	2.54%	2.81%	2.61%	2.54%	2.81%
1996	2.73%	2.65%	2.84%	2.90%	2.89%	2.75%	2.95%	2.88%	3.00%	2.99%	3.26%	3.32%	2.93%
1997	3.04%	3.03%	2.76%	2.50%	2.23%	2.30%	2.23%	2.23%	2.15%	2.08%	1.83%	1.70%	2.34%
1998	1.57%	1.44%	1.37%	1.44%	1.69%	1.68%	1.68%	1.62%	1.49%	1.49%	1.55%	1.61%	1.55%
1999	1.67%	1.61%	1.73%	2.28%	2.09%	1.96%	2.14%	2.26%	2.63%	2.56%	2.62%	2.68%	2.19%
2000	2.74%	3.22%	3.76%	3.07%	3.19%	3.73%	3.66%	3.41%	3.45%	3.45%	3.45%	3.39%	3.38%
2001	3.73%	3.53%	2.92%	3.27%	3.62%	3.25%	2.72%	2.72%	2.65%	2.13%	1.90%	1.55%	2.83%
2002	1.14%	1.14%	1.48%	1.64%	1.18%	1.07%	1.46%	1.80%	1.51%	2.03%	2.20%	2.38%	1.59%
2003	2.60%	2.98%	3.02%	2.22%	2.06%	2.11%	2.11%	2.16%	2.32%	2.04%	1.77%	1.88%	2.27%
2004	1.93%	1.69%	1.74%	2.29%	3.05%	3.27%	2.99%	2.65%	2.54%	3.19%	3.52%	3.26%	2.68%
2005	2.97%	3.01%	3.15%	3.51%	2.80%	2.53%	3.17%	3.64%	4.69%	4.35%	3.46%	3.42%	3.39%
2006	3.99%	3.60%	3.36%	3.55%	4.17%	4.32%	4.15%	3.82%	2.06%	1.31%	1.97%	2.54%	3.24%
2007	2.08%	2.42%	2.78%	2.57%	2.69%	2.69%	2.36%	1.97%	2.76%	3.54%	4.31%	4.08%	2.85%
2008	4.28%	4.03%	3.98%	3.94%	4.18%	5.02%	5.60%	5.37%	4.94%	3.66%	1.07%	0.09%	3.85%
2009	0.03%	0.24%	-0.38%	-0.74%	-1.28%	-1.43%	-2.10%	-1.48%	-1.29%	-0.18%	1.84%	2.72%	-0.34%
2010	2.63%	2.14%	2.31%	2.24%	2.02%	1.05%	1.24%	1.15%	1.14%	1.17%	1.14%	1.50%	1.64%
2011	1.63%	2.11%	2.68%	3.16%	3.57%	3.56%	3.63%	3.77%	3.87%	3.53%	3.39%	2.96%	3.16%

As described in the main report, the *Annual* inflation rate for a given calendar year is simply the arithmetic average of the inflation rates under each month column. This is because inflation rate in each month column is *not* the inflation over the duration of that month. It is the inflation between that month and the same month of the preceding year. Therefore, only the *Annual* inflation rate was used to inflate new and used frigate prices to 1 January 2012.

For example, consider the USS Estocin, which was deemed to have been new in 10 January 1981 and sold to Turkey 3 March 2003. The price when new was $P^{\mathrm{BY}}_{\mathrm{New}} = \$94,930,000$ in budget year USD, while the price when it was sold used was $P^{\mathrm{BY}}_{\mathrm{Used}} = \$21,021,143$ in budget year USD. The age at the time of sale was 22.2 years. To inflate both prices to 1 January 2012, the effective starting dates for new and used were rounded to 1 January 1981 and 1 January 2003, respectively.

The inflation for $P_{\text{New}}^{\text{BY}}$ was calculated by selecting the annual inflation percentages for 1981 to 2011:

$$\begin{aligned}
i_{\text{New}} \quad = \quad & \left(1+\frac{10.35\%}{100\%}\right)\left(1+\frac{6.16\%}{100\%}\right)\left(1+\frac{3.22\%}{100\%}\right)\left(1+\frac{4.30\%}{100\%}\right) \\
& \times\left(1+\frac{3.55\%}{100\%}\right)\left(1+\frac{1.91\%}{100\%}\right)\left(1+\frac{3.66\%}{100\%}\right)\left(1+\frac{4.08\%}{100\%}\right) \\
& \times\left(1+\frac{4.83\%}{100\%}\right)\left(1+\frac{5.39\%}{100\%}\right)\left(1+\frac{4.25\%}{100\%}\right)\left(1+\frac{3.03\%}{100\%}\right) \\
& \times\left(1+\frac{2.96\%}{100\%}\right)\left(1+\frac{2.61\%}{100\%}\right)\left(1+\frac{2.81\%}{100\%}\right)\left(1+\frac{2.93\%}{100\%}\right) \\
& \times\left(1+\frac{2.34\%}{100\%}\right)\left(1+\frac{1.55\%}{100\%}\right)\left(1+\frac{2.19\%}{100\%}\right)\left(1+\frac{3.38\%}{100\%}\right) \\
& \times\left(1+\frac{2.83\%}{100\%}\right)\left(1+\frac{1.59\%}{100\%}\right)\left(1+\frac{2.27\%}{100\%}\right)\left(1+\frac{2.68\%}{100\%}\right) \\
& \times\left(1+\frac{3.39\%}{100\%}\right)\left(1+\frac{3.24\%}{100\%}\right)\left(1+\frac{2.85\%}{100\%}\right)\left(1+\frac{3.85\%}{100\%}\right) \\
& \qquad \times\left(1+\frac{-0.34\%}{100\%}\right)\left(1+\frac{1.64\%}{100\%}\right)\left(1+\frac{3.16\%}{100\%}\right) \times 100\% - 100\% \\
= \quad & 173\%
\end{aligned}$$

The inflation for $P_{\text{Used}}^{\text{BY}}$ was calculated by selecting the annual inflation percentages for 2003 to 2011:

$$\begin{aligned}
i_{\text{Used}} \quad = \quad & \left(1+\frac{2.27\%}{100\%}\right)\left(1+\frac{2.68\%}{100\%}\right)\left(1+\frac{3.39\%}{100\%}\right)\left(1+\frac{3.24\%}{100\%}\right)\left(1+\frac{2.85\%}{100\%}\right) \\
& \quad \times\left(1+\frac{3.85\%}{100\%}\right)\left(1+\frac{-0.34\%}{100\%}\right)\left(1+\frac{1.64\%}{100\%}\right)\left(1+\frac{3.16\%}{100\%}\right) \times 100\% - 100\% \\
= \quad & 25\%
\end{aligned}$$

The new and used prices in terms of 2012 USD were then calculated:

$$P_{\text{New}}^{2012} = \$94,930,000 \times [100\% + (i_{\text{New}} = 173\%)]/100\% = \text{US}\$259,158,900$$

$$P_{\text{Used}}^{2012} = \$21,021,143 \times [100\% + (i_{\text{Used}} = 25\%)]/100\% = \text{US}\$26,276,429$$

Annex B: Confidence intervals for linear regression

The following confidence interval formulas for linear regression are taken from reference [39].

Consider a set of points $\{(x_i, y_i) \mid i = 1...n\}$ with an estimated linear relationship $y = \hat{\alpha} + \hat{\beta}x$. The standard error of the slope estimator $\hat{\beta}$ is calculated as

$$s_{\hat{\beta}} = \sqrt{\frac{\frac{1}{n-2}\sum_{i=1}^{n}\hat{\varepsilon}_i^{\,2}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \ ,$$

where $\hat{\varepsilon}_i$ is the residual for point i.

Corresponding to a significance level $\gamma \in [0, 1]$, the $(1 - \gamma)$ confidence interval for $\hat{\beta}$ is then

$$\left[\, \hat{\beta} - s_{\hat{\beta}} t^{*}_{n-2},\ \hat{\beta} + s_{\hat{\beta}} t^{*}_{n-2} \,\right]$$

where t^{*}_{n-2} is the $(1 - \gamma)^{\text{th}}$ quantile of the t distribution with $n - 2$ degrees of freedom. The standard error of the intercept estimator $\hat{\alpha}$ can then be calculated as

$$s_{\hat{\alpha}} = s_{\hat{\beta}} \sqrt{\tfrac{1}{n}\sum_{i=1}^{n} x_i^2} \ .$$

The $(1 - \gamma)$ confidence interval for $\hat{\alpha}$ is then

$$\left[\, \hat{\alpha} - s_{\hat{\alpha}} t^{*}_{n-2},\ \hat{\alpha} + s_{\hat{\alpha}} t^{*}_{n-2} \,\right] \ .$$

This page intentionally left blank.

Annex C: Normality testing

This annex details the normal probability plot and the Anderson-Darling test for the results presented in Section 3.3. The techniques described are established in the literature. For the benefit of analysts for whom these may be new, they are introduced with an emphasis on intuition and a practical overview. Notwithstanding the more advanced tests based on quantitative thresholds, the simpler Q-Q plot is important because of its illustrative nature, and because it is good practice to always graphically examine the data as a complement to quantitative threshold testing [22–25]. In particular, the use of confidence bands in the normal probability plot can intuitively convey a sense of confidence to sponsors (and to analysts). For small sample sizes, however, the effects of the discrete nature of the data points becomes pronounced and a more rigorous approach than the common practice is needed to generate the confidence bands. This is deferred to the next annex in order to preserve the purpose of this annex, which is to convey intuition; the plot presented in this annex is the "quick and dirty" version that is the prevailing norm, shown to illustrate its interpretation.

C.1 Normal probability plot

A normal probability plot [40] provides a simple visual check of the normality of a data sample, which in our case is the set of residuals from linear regression in the log domain. It is a special case of a Q-Q plot [41]. The idea is to plot the sorted set of n data against the n-quantiles of the hypothesized distribution (the normal distribution here). Notwithstanding sampling noise, a very linear plot would indicate a good correspondence with a normal distribution.

The Q-Q plot has a number of expedient features. For many distributions, including the normal distribution, the slope and offset of the plot reflect the mean and scaling differences between the data and the hypothesized distribution. If the hypothesized distribution is taken to be standard normal, the mean and standard deviation for the data can be read directly from the offset and slope. The vertical axis is usually used for the data, in which case left skew, right skew, and negative kurtosis show in the plot up as gross concavity, convexity, and as an "S" shape[35], respectively.

Conceptually, it is straightforward to determine the n-quantiles against which the data are plotted. Consider the random variable (RV) associated with the hypothesized PDF[36]. It is well known that the CDF[36] provides a monotonic transformation between the hypothesized RV and an RV with the standard uniform distribution $U(0, 1)$ [42–45][37]. So the interval [0, 1] can be partitioned into n equal bins whose boundaries can be reverse transformed into the n-quantiles of the hypothesized PDF.

While this provides the intuitive basis for Q-Q plots, the details have "occasioned much discussion" [41]. A main point of divergence concerns the pairing of each of the sorted data with the corresponding n-quantile of the hypothesized PDF. The problem is that, by definition, quantiles mark the *boundaries* of the bins to be matched with each datum. In order to match the data with the *centres* of the bins, various adjustments to the boundaries have been used. As an alternative to the concept of

[35] Convexity in the lower regime and concavity in the upper regime.

[36] PDF = Probability density function. CDF = Cumulative distribution function.

[37] This transformation is typically used, for example, in generating random data with a specific distribution in Monte Carlo simulations.

quantiles to guide this method, a common and less arbitrary approach is to base the method on *order statistics* [46].

For an n-size random sample of our hypothesized distribution, the k^{th} order statistic, often denoted $x_{(k)}$, refers to the k^{th} largest datum. Over repeated n-size samples, $x_{(k)}$ takes on its own distribution. In principle, this distribution can be determined from the hypothesized normal distribution for our data set. It is the median of the distribution for $x_{(k)}$ which our k^{th} largest residual is plotted against. These order statistic medians are approximated by using the inverse of the hypothesized CDF to transform the order statistic medians of $U(0, 1)$, similar to what was done above for quantiles. In turn, the order statistic medians for $U(0, 1)$ are approximated by

$$\begin{array}{c} U(0,1) \text{ order} \\ \text{statistic median} \end{array} = \begin{cases} 1 - 0.5^{1/n}, & i = 1 \\ \dfrac{i - 0.3175}{n + 0.365}, & i = 2, 3, \ldots, n - 1 \\ 0.5^{1/n}, & i = n \end{cases}$$

Figure C.1 plots the residuals against the $N(0, 1)$ order statistic medians generated from the above $U(0, 1)$ order statistic medians. The relatively linear plot has a coefficient of determination $r^2 = 95\%$ and contains none of the features that would suggest skewness or excess kurtosis[38]. The slope clearly reflects the tabulated standard deviation of the residuals, 0.144. The data points are sorted, so adjacent points are not independent. Hence, runs of consecutive data points above and below the line of best fit are not unexpected [47, 48].

Figure C.2 shows approximate 95% confidence bands to indicate how well the residual quantiles match those of a normal distribution. For simplicity, both axes are scaled according to the standard deviation of the residuals. Each data point corresponds to a residual, and sample CDF would completely match a normal CDF if the data points all fell on the $y = x$ line (not the line of best fit). For realistic data, which wanders off the straight line, the horizontal distances between each point and the 95% bands represent the maximum permissible horizontal distances between the data points and the straight line. Therefore, if the confidence bands fail to straddle the straight line at any point, the null hypothesis that the residuals are normal is rejected at a 5% significance level. In Figure C.2, the straight line is fully contained within the 95% bands, so there is no appreciable evidence against normality.

The generation of confidence bands is described in Annex D. The confidence bands in Figure C.2 result from the prevailing approach found in the literature, which evaluates the confidence intervals only at the sample data points, then interpolates between them. While this is accurate for large sample sizes, Annex D describes the ambiguities that arose for the small sample size of this report. The ambiguities centre around the selection of quantiles and lead to differences in results that are more prominent for small sample sizes (a likely situation when data is hard to obtain). To resolve this, the strict definition of the test statistic is used to avoid the ambiguity, and more rigorous confidence

[38] The normal distribution has a kurtosis of three. Excess kurtosis refers to how much the kurtosis of the data differs from three and is zero for a normal distribution.

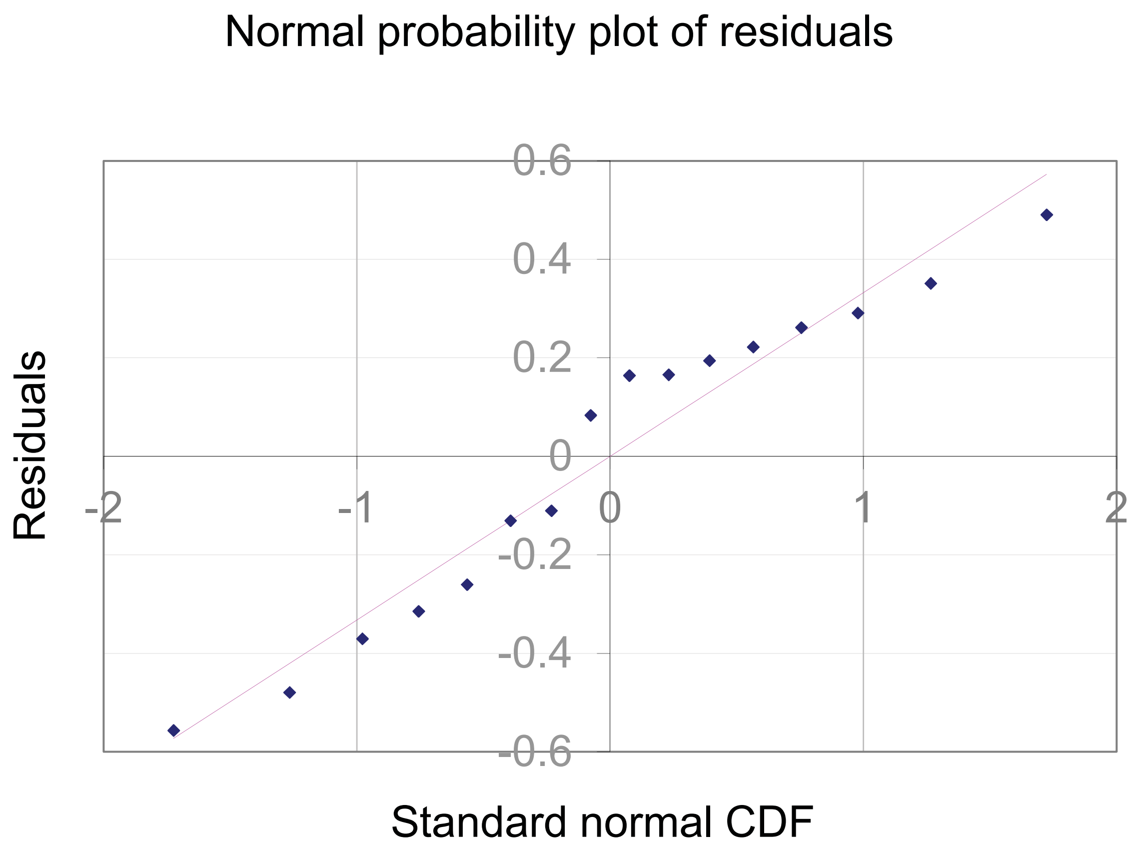

Figure C.1: *Normal probability plot of residuals.*

bands are derived from first principles. This yields a more conservative confidence bands in the sense that they can be breached when they otherwise would not when using ad-hoc quantiles and linear interpolation. It is the Q-Q plot in Annex D that is presented in the body of the report.

C.2 Anderson-Darling test

The Anderson-Darling test [49] compares the empirical CDF for the data with the CDF for the hypothesized distribution. It performs well for small sample sizes. It is similar to the Cramér-von Mises criterion [50] in that it integrates the square of the differences between the two CDFs. However, it weights the differences in the tails of the distributions more heavily than the centre. Since the CDF transforms data from a given distribution to $U(0,1)$, the test can be seen as comparing the data's order statistics in the $U(0,1)$ domain. The comparison is between the data transformed to $U(0,1)$ using the hypothesized PDF and the empirical CDF. For a sample size n, the test statistic is

$$A^{*2} = \left(\frac{25}{n^2} - \frac{4}{n} - 1\right)\left(n + \frac{1}{n}\sum_{i=1}^{n}\left\{(2i-1)\ln\Phi(Y_i) + \left[2(n-i)+1\right]\ln\left[1-\Phi(Y_i)\right]\right\}\right)$$

where the Y_i's are the sorted z-scores of the data and $\Phi(\cdot)$ is the standard normal CDF. We reject the null hypothesis that the two CDFs are the same if A^{*2} exceeds a threshold that depends on: (i) the desired statistical significance and (ii) how many of the hypothesized distribution parameters are

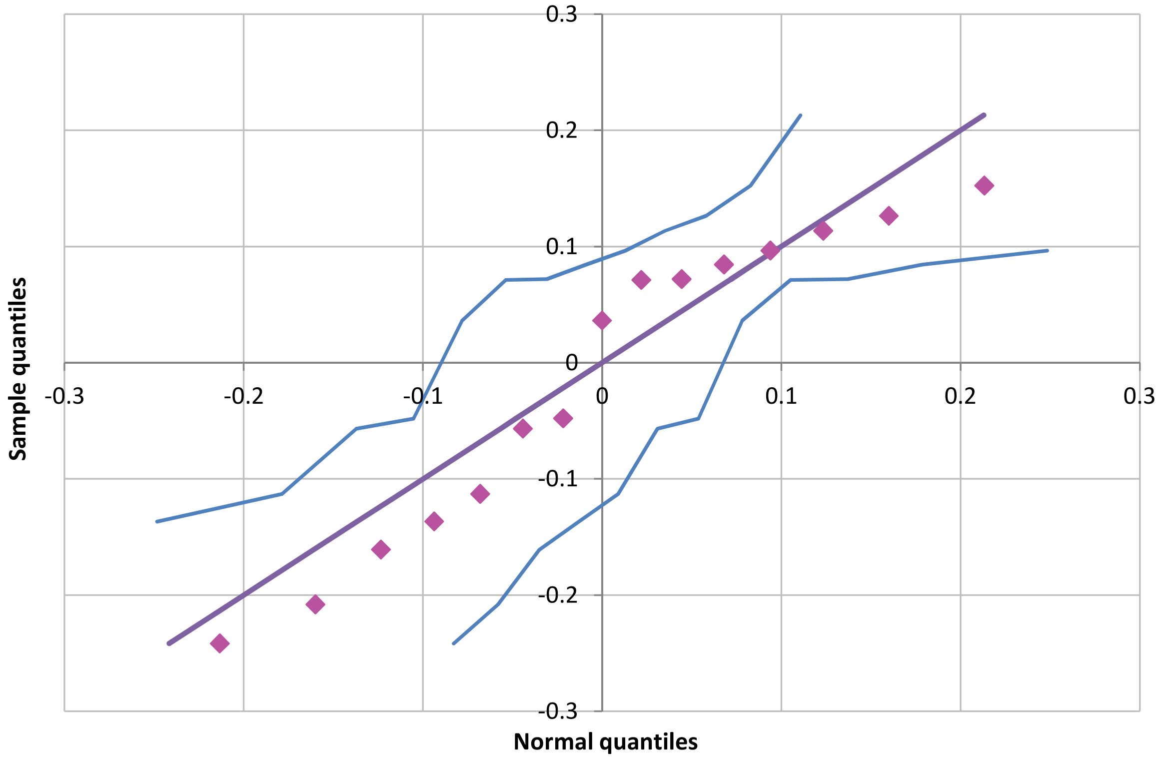

Figure C.2: *Approximate 95% confidence bands.*

asserted as opposed to estimated from the sample. In the normal PDF hypothesized here, both the mean and variance parameters are unknown and estimated from the data.

Over the past half century, various adjustments to the formulas for A^{*2} have been developed, along with tables of thresholds for rejection of the hypothesized distribution at various significance levels[39]. Table C.1 compiles thresholds for a normal distribution.

For our residuals, $A^{*2} = 0.516$ exceeds none of the tabulated thresholds, indicating that the null hypothesis of normally distributed residuals cannot be rejected. Note that the 1986 thresholds (second row) use an adjusted A^{*2} formula. For our residuals, this yields an adjusted $A^{*2} = 0.473$, which still exceeds none of thresholds. Hence, the p-value is decidedly greater than 0.15. In terms of area under the PDF for a test statistic *when* H_0 *is true*, the p-value is the portion of the tail beyond the value corresponding to the data sample. Small p-values are further out the tail, implying that H_0 is less plausible. Weak evidence against H_0 is typically considered to emerge when p-values drop to 0.1, getting stronger with lower p-values. Therefore, the Anderson-Darling test yields no appreciable evidence against null hypothesis that the residuals are normally distributed.

In determining whether tabulated Anderson-Darling thresholds were appropriate for our sample size of 16, we found that it necessary to go beyond a superficial information search. The challenge

[39] Thresholds for Anderson-Darling and other statistics are obtained using a variety of ways: Typically, analytically for large sample sizes, and using Monte Carlo simulation with numerical fitting for series of small sample sizes.

Table C.1: *Anderson-Darling test statistic thresholds for the hypothesis of a normal distribution. For any threshold exceeded by A^{*2}, H_0 (normality) is rejected with the associated significance level.*

Source	Significance						Minimum sample size
	15%	10%	5%	2.5%	1%	0.5%	
Threshold 1974 [49]	0.576	0.656	0.787	0.918	1.092		∼10
Threshold 1986 [51][a]	0.561	0.631	0.752	0.873	1.035	1.159	8
Threshold [43]		0.632	0.751	0.870	1.029		Unspecified

[a] Chapters 4 and 9.

has to do with the empirical nature of the methods, which is the very aspect that allows the tests to be characterized for small sample sizes. Such studies typically examine a variety of test statistics, suites of sample sizes, and a number of different scenarios in which different amounts of information are known about the PDF with which the data is being compared. As an example of the latter, the scenario most frequently encountered by analysts is the need to compare data against a normal PDF with unknown mean and variance [52], i.e., the parameters are determined from the data, as they are in this report. Though the general message is that very small sample sizes can be characterized with these approaches, the studies cross so many facets that it can be difficult to definitively determine that a certain minimum sample size applies specifically to the Anderson-Darling statistic, with a certain curve fitting, and specifically for a normal PDF with unknown parameters. At the time of writing, even a relatively high quality source of open source information like Wikipedia, which is subject to world-wide scrutiny, over-generalized the minimum sample size of 5 for the posted Anderson Darling thresholds[40]. It is highly advisable, therefore, that original papers be consulted for thresholds. For this report, these seem to be references [52, 53]. Reference [52] is not entirely unambiguous, but strongly suggests a minimum sample size of 10. Reference [53] corroborates this with its characterization of the curve fitting, though it does not examine Anderson-Darling specifically. Fortunately, one source [51] is unequivocal about its minimum size of 8.

[40]It applies to normal PDFs whose means (μ's) are known.

This page intentionally left blank.

Annex D: Confidence bands in normal probability plots

Confidence bands are a useful visualizing method for conveying how much one can trust a Q-Q plot. It can give an indication of whether greater effort is needed in checking normality. Confidence bands are based on the Kolmogorov-Smirnov (KS) test of equality between two CDFs [54–56]. This chapter describes the intuition behind the statistics used, along with an ambiguity that was encountered in how the quantile points are chosen. The ambiguity matters more for small sample sizes. In order to improve the confidence with which this visualization can be used in the future, the cause of the ambiguity is graphically explored. A simple graphical solution is derived from first principles. Figure C.2 is duplicated here as Figure D.1 for ease of comparison.

D.1 Confidence bands from the Kolmogorov-Smirnov test

Often, the CDF for a data sample, CDF_{Sample}, is being tested against a hypothesized normal CDF, CDF_{H0}. That is, the data sample is deemed to be the result of sampling a random variable, and H_0 is the hypothesis that CDF_{Sample} is the same as CDF_{H0}. The empirical CDF_{Sample} is defined as one might expect:

$$CDF_{\text{Sample}}(x) \;=\; \left\{ \text{Portion of the data sample } \;\leq\; x \right\} \tag{D.1}$$

The KS statistic of similarity between the CDF_{Sample} and CDF_{H0} is simply the maximum difference between the two CDFs over the value range of the random variable, which we will refer to as $|\Delta CDF|_{\text{max}}$. For the case that H_0 is true, repeated data samples lead to a distribution for $|\Delta CDF|_{\text{max}}$, allowing thresholds to be defined for various levels of significance. These thresholds are tabulated; for a given sample size, specifying a significance level automatically determines the maximum possible difference between CDF_{H0} and CDF_{Sample} in order to avoid rejection of H_0. Note that, at a given confidence level, the thresholds for $|\Delta CDF|_{\text{max}}$ were determined for the case that CDF_{H0} is within distance $|\Delta CDF|_{\text{max}}$ of CDF_{Sample} at *all* data points.

Imagine that we have looked up our required threshold value for $|\Delta CDF|_{\text{max}}$ for a sample size of 16 and significance $\alpha = 5\%$. Let us refer to this threshold value as $|\Delta CDF|^{\text{thr}}_{\text{max}}$. In coming up with the confidence bounds, the path of CDF_{Sample} is taken to be the reference curve, and

$$CDF_{\text{Sample}} \pm |\Delta CDF|^{\text{thr}}_{\text{max}}$$

then defines upper and lower bounds on the value for CDF_{H0} in order to avoid rejection of H_0. Hence, these bounds must straddle the CDF_{H0} curve. To have these bounds show up on the Q-Q plot, they are transformed back into the domain of the random variable (which is the domain of the Q-Q plot) by applying the reverse transform CDF^{-1}. In this domain, the bounds must still straddle the hypothesized distribution because $CDF(\cdot)$ and $CDF^{-1}(\cdot)$ are monotonic and hence preserve the inequalities that define the bounds. However, the hypothesized distribution is just a straight line on the Q-Q plot, so the bounds must straddle the straight line. Note that, even though the confidence bands are offset from CDF_{Sample} by a constant $|\Delta CDF|^{\text{thr}}_{\text{max}}$, they do not track the Q-Q plot with

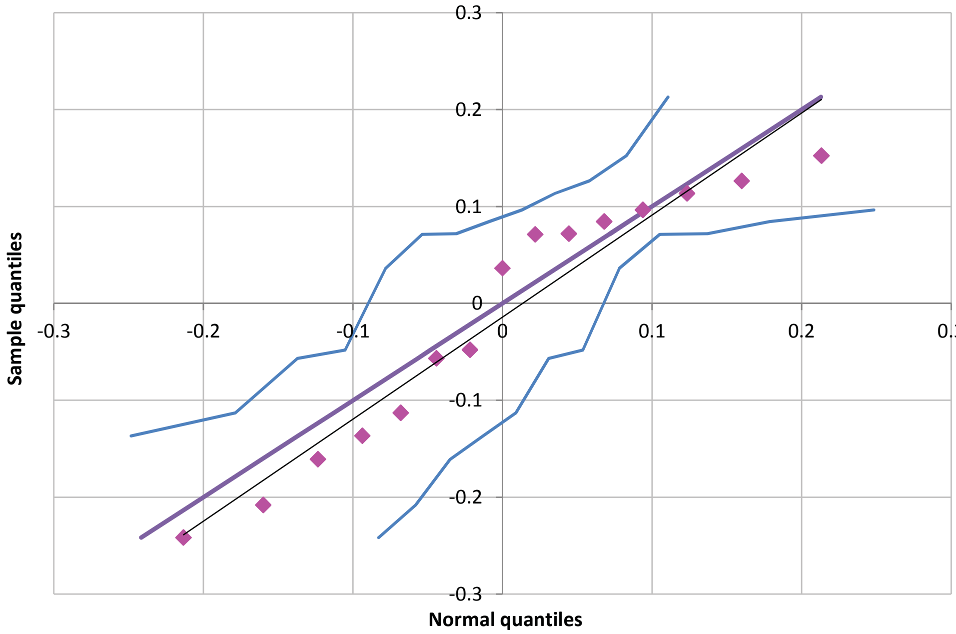

Figure D.1: *95% confidence bands on Q-Q plot. The line of best fit is shown in black.*

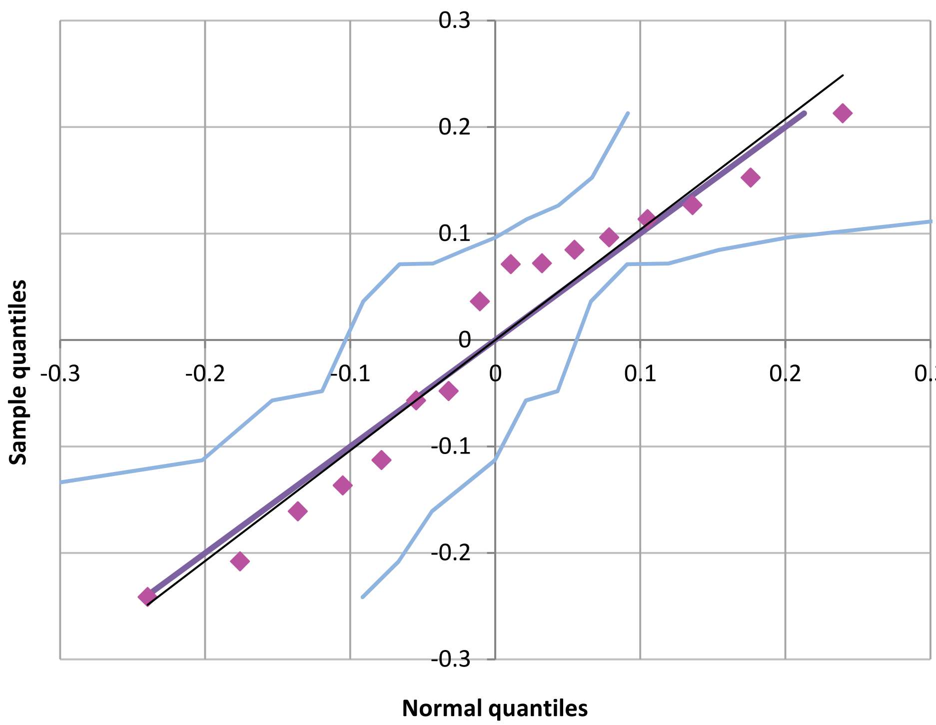

Figure D.2: *Q-Q plot and 95% confidence bands, using Filliben's estimates of U(0,1) order statistic medians for the plotting positions. The line of best fit is shown in black.*

constant horizontal offset in Figure D.1 because the reverse transformation Φ^{-1} from the $U(0,1)$ domain is nonlinear.

D.2 Lilliefors test when distribution parameters unknown

While the above describes how confidence bands are determined on a Q-Q plot, the original KS test is restricted in its practical application because the thresholds are for the situation in which the parameters of CDF_{H0} are known. If instead the parameters are estimated from the sample data, then CDF_{Sample} is closer to CDF_{H0} and $|\Delta CDF|^{\mathrm{thr}}_{\mathrm{max}}$ is lower. The KS thresholds can be significantly misleading. Thresholds for the case in which the CDF_{H0} parameters are estimated from the sample data have been determined through Monte Carlo analysis [57,58]. Note that the calculation of the test statistic $|\Delta CDF|_{\mathrm{max}}$ does not change, but the thresholds for the different confidence levels do. The use of the adjusted thresholds is referred to as the Lilliefors test, after its developer. It is the $\alpha = 5\%$ value of this adjusted threshold for $|\Delta CDF|^{\mathrm{thr}}_{\mathrm{max}}$ that was used to determine the 95% confidence bands in Figure D.1.

D.3 Asymmetry from classical quantile boundaries

In generating confidence bands, some anomalies were encountered with the above definition of CDF_{Sample}. In Section C.1, for the Q-Q plot without confidence bands, the data points associate the sorted sample data with the medians of the order statistics, but there are other options. The exact locations chosen for evaluating CDF_{H0} are known as *plotting positions*. For the order statistic medians in the $U(0,1)$ domain, the approximations presented in Section C.1 are known as *Filliben's estimate* [41]. These yield a symmetric set of quantiles both in the random variable and the U(0,1) domain, which the CDF transforms the data into. In contrast, the quantiles for the empirical CDF_{Sample} as defined in equation D.1 is quite asymmetric – the problem has similarities to the asymmetric offset in quantile bins described in Section C.1. In the $U(0,1)$ domain, the first quantile is $1/N_{\mathrm{Points}}$, well away from the bottom, while last ($N^{\mathrm{th}}_{\mathrm{Points}}$) quantile is 1 i.e., at the top end. These are the plotting positions used in the Q-Q plot with confidence bands in the Section C.1, and in Figure D.1.

This asymmetry is systematic, and leads to systematic asymmetry in the normal distribution quantiles, against which the data (the residuals) is plotted. Sampling noise notwithstanding, the residuals are not expected to have such an asymmetry unless there is something seriously wrong with the analysis. For comparison, Figure D.2 shows the resulting Q-Q plot using Filliben's estimate for symmetric plotting positions. Comparing Figures D.1 and D.2 reveals that the uppermost point cannot be plotted using the asymmetric plotting positions of Figure D.1 because the $N^{\mathrm{th}}_{\mathrm{Points}}$ quantile is 1, which the inverse transform $\Phi^{-1}(1)$ maps to ∞. Note also that the line of best fit does not pass through the origin in Figure D.1, due at least in part to the lost data point. Finally, compared to Figure D.2, the remaining data points are also shifted to the right because the quantiles are biased upward. It should be clear this half-quantile shift has a more significant impact when the sample size is small.

Figures D.1 and D.2 illustrate the strong intuitive reason for using Filliben's estimate for the $U(0,1)$ quantiles, in contrast to the asymmetric quantiles that result from the above definition of CDF_{Sample}. However, there is a strong theoretical premise for the latter. The important inequalities that lead to the straddling of the normal plot by the confidence bounds follow directly from the Kolmogorov-Smirnov statistic, which is based on the empirical CDF_{Sample} as defined above. Thus it appears to be necessary to have the upwardly biased quantiles, despite the un-intuitiveness of this.

For this study, both of the Q-Q plots are within the 95% confidence interval for nonrejection of normality, so the difference seems immaterial. However, the ambiguity surrounding the plotting position needs to be resolved for the proper use of confidence bands in the future. A perusal of discussion on plotting positions in the literature did not offer a clear answer [59, 60]. The following sections delve into the details of confidence bands from first principles to show that neither of the confidence bands in Figures D.1 and D.2 are strictly correct for small sample sizes. In fact, the inaccuracies of conventional confidence bands go beyond simply choosing plotting positions. We found that, rather than linearly interpolating between discrete quantile values, the quantiles have to be determined from the continuous CDFs.

D.4 Confidence bands from a dense sample CDF

To get some insight into the nature of the ambiguity, Figure D.3 shows the "dense" CDFs CDF_{H0} and CDF_{Sample}, i.e., evaluated not just at the data sample values. While the empirical CDF_{Sample} is typically defined as in equation D.1, this section empirically explores how ΔCDF behaves in between the data points. CDF_{Sample} has a staircase shape because at every x location where there is a data point, CDF_{Sample} increments by $1/N_{\text{Points}}$. In contrast to the aforementioned practice of taking $CDF_{\text{Sample}} \pm |\Delta CDF|^{\text{thr}}_{\text{max}}$ as the bounds on CDF_{H0}, we initially take $CDF_{H0} \pm |\Delta CDF|^{\text{thr}}_{\text{max}}$ as the bounds on CDF_{Sample} because the bounds are less noisy and easier to follow. All four curves in Figure D.3 are plotted against CDF_{H0} in Figure D.4. This is a parametric plot with x (the domain of the random variables) as the free parameter. The μ is zero because it is simply the mean of the residuals, while σ is the residuals' sample standard deviation.

An important feature in Figures D.3 and D.4 that is missing from the previous Q-Q plots with confidence bands are the black bars in the middle, which are just marginally inside the bounds – this lack of a "safety margin" is visible only if the entire CDF_{Sample} is considered rather than just the point values in the data sample. In fact, it should be clear that CDF_{Sample} could breach the confidence bands even if all the data samples reside inside the confidence bands, and that a plot of the dense CDF_{Sample} is needed to see this. For small sample sizes, therefore, it is important to check all the corner points of the "staircase" CDF, including the open and closed ends of each piecewise uniform interval. In effect, the plot involves $2N_{\text{Points}}$[41].

We eventually want to depict boundaries that are constant offsets from CDF_{Sample}, entirely containing CDF_{H0} (or not, as the case may be). Therefore, we next take CDF_{Sample} as the reference and $CDF_{\text{Sample}} \pm |\Delta CDF|^{\text{thr}}_{\text{max}}$ as the upper and lower bounds for CDF_{H0}. Figure D.5 is a modification of Figure D.4 to show this. The points of marginal containment are the same as in Figure D.4 (0.6 and 0.7 along horizontal axis) except that the CDF_{Sample}-based boundaries descend to meet CDF_{H0} rather than the CDF_{Sample} descending to meet the CDF_{H0}-based boundaries.

[41] There are doubtlessly many ways to prepare the data for this, but the simplest is to scan $CDF_{\text{Sample}}(x)$ from left to right, allocating a data point the beginning and ending of each line segment. There will be 2 data points with the same x-value and 2 data points with the same y-value. For each of CDF_{Sample} and $CDF_{\text{Sample}} \pm |\Delta CDF|^{\text{thr}}_{\text{max}}$, there will be twice as

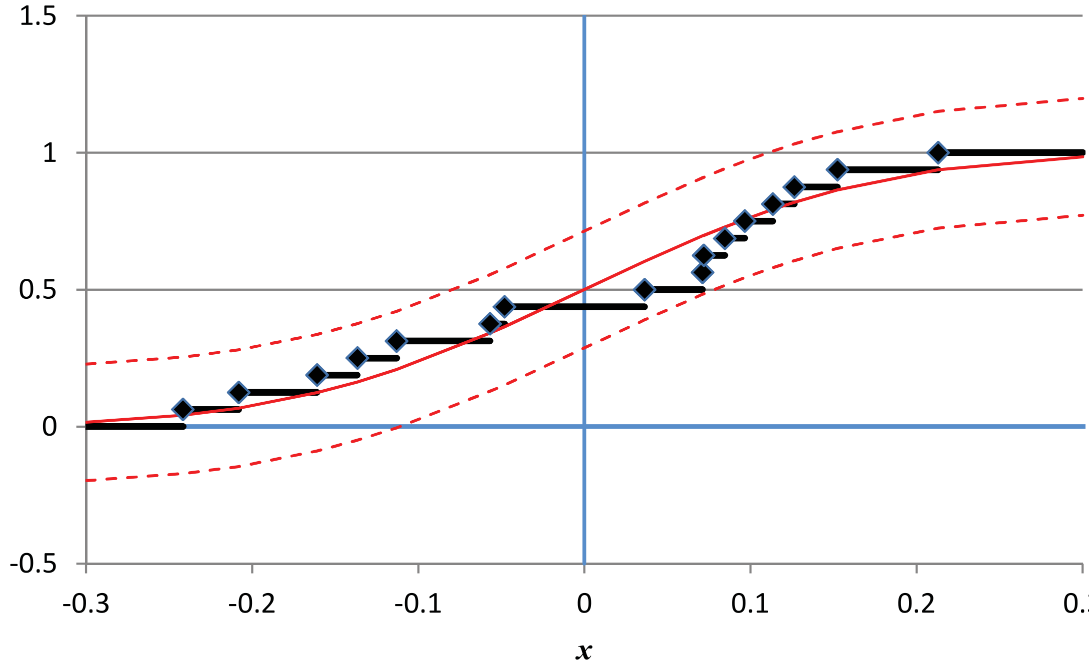

Figure D.3: *CDFs for H_0 (CDF_{H0}, solid red) and the data sample (CDF_{Sample}, black). The variable x is the domain of the random variables. The dashed lines are the 95% confidence bounds at $CDF_{H0} \pm |\Delta CDF|^{thr}_{max}$.*

Note that the confidence bands in Figure D.5 are drawn continuously rather than as vertically displaced copies of CDF_{Sample}. This is for ease of interpretation. Strictly speaking, the upper and lower bounds are exact copies of CDF_{Sample}, vertically displaced by $\pm|\Delta CDF|^{\text{thr}}_{\text{max}}$. However, that was found to be visually messy and the points at which there is a lack of margin between the bounds and CDF_{H0} was not immediately clear. Keeping in mind that only the upper-left corners of the steps in the staircase plot of $CDF_{\text{Sample}} \pm |\Delta CDF|^{\text{thr}}_{\text{max}}$ correspond to data samples, it should again be clear from Figure D.5 that, as in Figure D.4, the upper and lower bounds may appear to be far away from CDF_{H0} if the comparison is only made at the locations of the data points. For small sample sizes, however, this is not necessarily an adequate test of being within the confidence band. Again, all the corner points of the steps in the staircase need to be tested.

D.5 Reverse transformation into domain of random variables

A final modification is required for confidence bands on Q-Q plots. Figures D.4 and D.5 plot CDFs against CDFs, i.e., in the $U(0,1)$ domain. These probability plots are a variation of Q-Q plots known as P-P plots. Each type of probability plot has its advantages. The advantages of Q-Q plots are briefly described in Section C.1, and these features allow a data set to be compared to an entire family of

many points as there are residuals. It may be helpful to keep track of whether each line ending is an open or closed end.

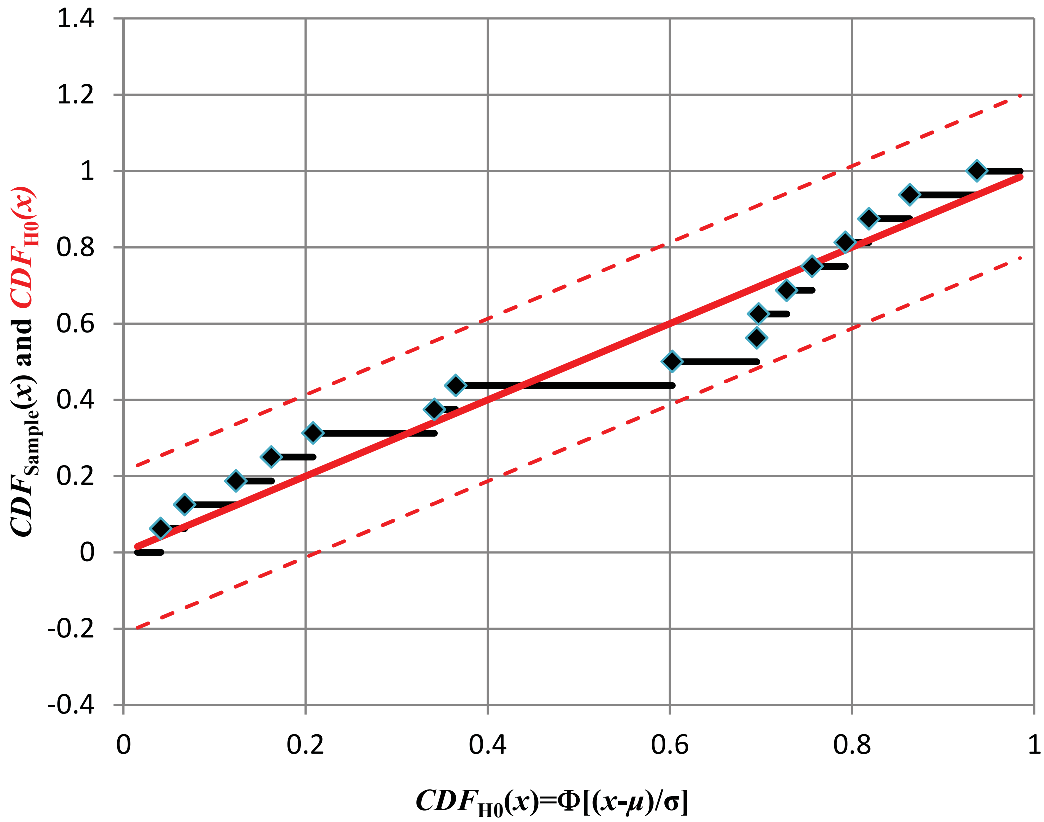

Figure D.4: *CDF_{Sample} and CDF_{H0} versus the latter, with 95% confidence bands $CDF_{H0} \pm |\Delta CDF|^{thr}_{max}$ flanking the latter (dashed lines).*

distributions. Generating Q-Q plottable data requires transforming the plots from $U(0, 1)$ domain to x, the domain of the random variables. First, write out the boundary relationship explicitly [55]:

$$CDF_{\text{Sample}}(x) - |\Delta CDF|^{\text{thr}}_{\text{max}} \;\leq\; CDF_{\text{H0}}(x) \;\leq\; CDF_{\text{Sample}}(x) + |\Delta CDF|^{\text{thr}}_{\text{max}}$$

Keeping in mind that the results apply to general shaped CDF_{H0}[42], let us specifically consider the normal $CDF_{\text{H0}}(x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right)$. Then:

$$\mu + \sigma\,\Phi^{-1}\left[CDF_{\text{Sample}}(x) - |\Delta CDF|^{\text{thr}}_{\text{max}}\right] \;\leq\; x \;\leq\; \mu + \sigma\,\Phi^{-1}\left[CDF_{\text{Sample}}(x) + |\Delta CDF|^{\text{thr}}_{\text{max}}\right]$$

The resulting confidence bands are shown in Figure D.6, along with the Q-Q plot using the dense CDF_{Sample}. This may seem unintuitive at first, so it is worthwhile going over exactly why the plot takes the form that it does. First, let us examine the Q-Q plot itself, followed by the confidence bands.

[42] Probability distributions are compared for more purposes than just checking normality.

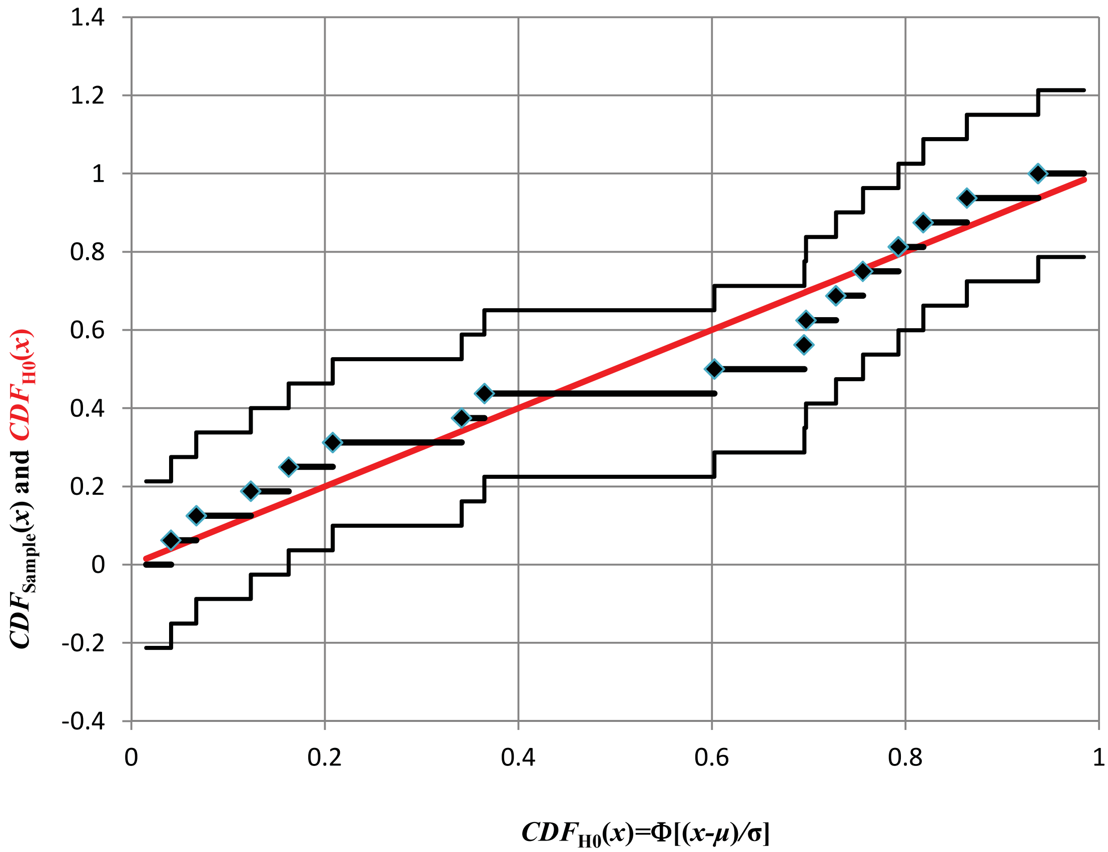

Figure D.5: *CDF_{Sample} and CDF_{H0} versus the latter, with 95% confidence bands $CDF_{Sample} \pm |\Delta CDF|^{thr}_{max}$ flanking the former (thin black boundaries).*

D.6 Making sense of the Q-Q plot

The points in Figure D.6 corresponding to the data sample values are in the same positions as those in Figure D.1, as one would expect. Since Figure D.6 doesn't just evaluate the Q-Q plot at the data sample values, however, there are stems emerging upward from each data point. To see why this makes sense, it necessary to view the Q-Q plot not just as scatter plot of N_{points}-quantiles for 2 distributions being compared. Instead, a "quantile" is often interpreted more liberally as an inverse CDF – that is, the p^{th} quantile of a distribution ($p \in [0, 1]$) is that value of a random variable X for which $P(x \leq X) = p$. Applying this to Figure D.3, a Q-Q plot is generated by a parametric variable p scanning the range $y : 0 \rightarrow 1$. At each value of p, this generates a matching pair of x values from $CDF_{\mathrm{H0}}(x) = p$ and $CDF_{\mathrm{Sample}}(x) = p$, to be used as coordinates in the Q-Q plot. Recall that the horizontal and vertical axes of the Q-Q plot are the quantiles of CDF_{H0} and CDF_{Sample}, respectively. For a given value of p, let's refer to the matching pair of x values as x_{H0} and x_{Sample}, respectively.

From Figure D.3, let us examine what is generated by $CDF_{\mathrm{H0}}(x_{\mathrm{H0}}) = p = CDF_{\mathrm{Sample}}(x_{\mathrm{Sample}})$ as $p : 0 \rightarrow 1$ and how this is reflected in the Q-Q plot of Figure D.6.

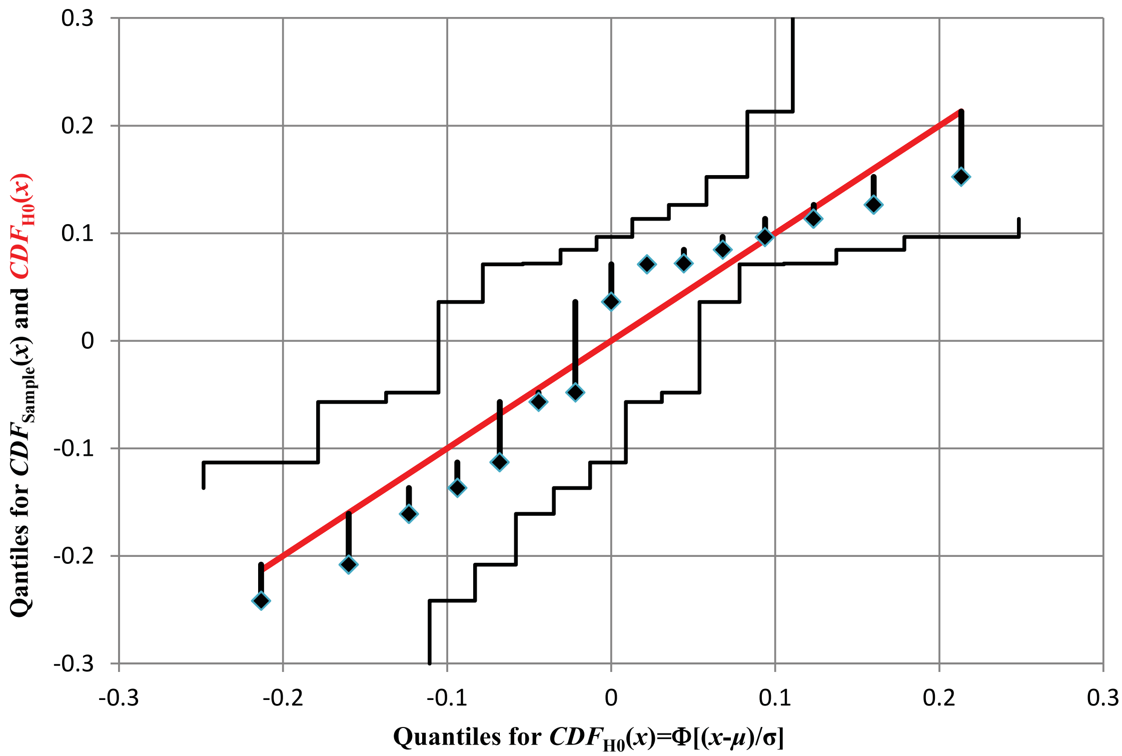

Figure D.6: *The dense Q-Q plot with continuous quantile and 95% confidence bands.*

1. For $p : 0 \to (1/N_{\text{Points}})^-$, $x_{\text{H0}} : -\infty \to -2.4$, but x_{Sample} exists only for $p = 0$ (where x_{H0} is $-\infty$). Hence, nothing can be plotted on the Q-Q plot for $p : 0 \to (1/N_{\text{Points}})^-$.
2. At $p = (1/N_{\text{Points}})$, x_{H0} takes on a single value ~ -0.215 and x_{Sample} takes on a range of values $[-0.240, -0.210)$. This is exactly what shows up on the Q-Q plot.
3. For $p : (1/N_{\text{Points}})^+ \to (2/N_{\text{Points}})^-$, x_{H0} takes on continuous finite values, but x_{Sample} does not exist. Nothing can be plotted on the Q-Q plot.
4. At $p = (2/N_{\text{Points}})$, x_{H0} takes on a single value ~ -0.155 and x_{Sample} takes on a range of values $[-0.210, -0.170)$. This is exactly what shows up on the Q-Q plot.

This is repeated until $p = N_{\text{Points}}/N_{\text{Points}} = 1$. However, the pattern that emerges in the above walk-through is clear – an alternation between nothing and vertical line segments with open tops and closed bottoms. This shows the following rules in the Q-Q plot of Figure D.6, which will be used to draw the confidence bands.

1. The line spectrum taken on by x_{H0} follows directly from discrete nature of CDF_{Sample}, even though CDF_{H0} is continuous. More specifically, it is the values of p at which the CDF_{Sample} lines exist that determine the values of x_{H0} where lines exist in the Q-Q plot, i.e., the discrete values for $p = CDF_{\text{Sample}}$ pick out the values at which to evaluate $x_{\text{H0}} = CDF^{-1}_{\text{H0}}(p)$.

2. The closed end of each line segment is at the lower end, as expected.

3. The x_{Sample} value ranges (vertical axis) of the line segments mirror their value ranges in the CDF plot of Figure D.3.

D.7 Making sense of the confidence bands

With the above understanding of how CDF_{Sample} influences the Q-Q plot, the explanation of the upper and lower confidence bands becomes straightforward. The upper band is simply an upward shift of CDF_{Sample} in Figure D.3. From rule #1 above, this increases the x_{H0} of the corresponding vertical line segment in Figure D.6, i.e., shifts it right. However, from rule #3, it does not change the x_{Sample} value range that it spans, i.e., it does not move up or down in Figure D.6. Similarly, the lower bound is a left-shifted version of the Q-Q plot. This is exactly what is shown in Figure D.6.

This more detailed manner of arriving at confidence bands is premised on the strict definitions of $CDF_{\text{Sample}}(x)$ (equation D.1) and the test statistic $\max_x \left|CDF_{\text{H0}}(x) - CDF_{\text{Sample}}(x)\right|$. It is a more pessimistic approach than the use of ad-hoc quantiles and the linear interpolation of confidence bands between the data points in Figures D.1 and D.2; as described, the staircase bands have corners that can cause a breach of the straight line of ideal fit when the linearly interpolated bands do not. This conservatism is suitable if an analyst wants to present a cautious (i.e., rigorous) view of the confidence bands for small sample sizes.

The reader may have noticed that in order for the more detailed confidence bands to be valid, the tables of $|\Delta CDF|^{\text{thr}}_{\text{max}}$ being consulted must have been generated in accordance with the strict definitions above, including all the corners of the staircases. The treatment in publications on empirical CDF testing can be quite terse [51, 61, 62], and it can be unclear whether the formulas for the test statistics account for the staircase corners of the CDFs[43]. Reference [63] provides a concise and lucid treatment that makes it clear that the corners are being accounted for. The explanation can be related to Figure D.3 as follows. Each sample data point x_i ($i = 1, 2, ..., n$) corresponds to a vertical step in the staircase where $|\Delta CDF|$ is evaluated both at the bottom and top of the step as $\left|CDF_{\text{H0}}(x_i^-) - CDF_{\text{Sample}}(x_i^-)\right|$ and $\left|CDF_{\text{H0}}(x_i) - CDF_{\text{Sample}}(x_i)\right|$. This simplifies to $\left|CDF_{\text{H0}}(x_i) - (i-1)/n\right|$ and $\left|CDF_{\text{H0}}(x_i) - i/n\right|$.

To recognize this in the literature, note that the maximization of the above is often presented as $D = \max(D^+, D^-)$, where

$$\begin{aligned} D^+ &= \max_i \left|i/n - z_{(i)}\right|, \\ D^- &= \max_i \left|z_{(i)} - (i-1)/n\right|, \end{aligned}$$

$z_{(i)}$ corresponds to $CDF_{\text{H0}}[x_{(i)}]$, where $x_{(i)}$ is the i^{th} order statistic. Our notation simply uses x_i because we start by referring to Figure D.3, where the samples are already sorted.

---

[43] Examples and illustrations such as Chapter 4 of [62] and Chapters 5-6 of [51] suggest that the corners are accounted for, but this is far from an explicit statement of the important fact.

This page intentionally left blank.

Annex E: Increasing real cost of new FFG-7 with time

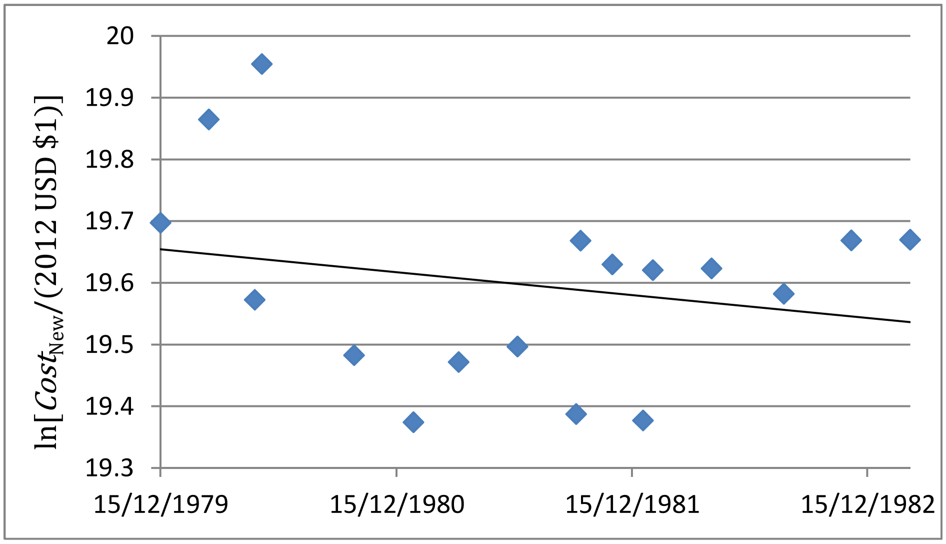

Figure E.1: *Log transformed cost of a new FFG-7 in 2012 USD as a function of its date of initial service. In this small sample of closely spaced data in time, it was not possible to detect a published growth trend in the real cost of platforms [38]. Nor was there evidence against such a growth trend.*

Table E.1: *Projected cost of "modern" FFG-7 built in 1992 [2012 USD], taken as the mean between air defence and ASW vessel classes. The "specific" cost [2006 British pounds (GBP) per ton] of building in 2006 and 2% real cost growth are first taken from [38]. The specific cost is converted to 2006 USD using the 2006 average of 0.5436 USD/GBP [64]. The 2% real cost growth is then applied as a deflation to get the specific cost of building in 1992 (denomination still in 2006 USD). U.S. inflation is then applied to get the 2012 USD specific cost of building in 1992. This is multiplied by the FFG-7 full-load displacement to get the 1992 cost per FFG-7 in 2012 USD.*

Vessel class	Lower quartile	Median	Upper quartile	Annual increase, real dollars
	Specific cost in ***2006*** [***2006 GBP*** /ton]			
Air defence vessel	70,000	93,000	110,000	2%
ASW vessel	84,000	100,000	140,000	2%
Mean	77,000	96,500	125,000	2%
	Specific cost in ***2006*** [***2006 USD*** /ton]			
Mean	41,857	52,457	67,950	2%
	Specific cost in ***1992*** [***2006 USD*** /ton]			
Mean	31,723	39,756	51,498	
	Specific cost in ***1992*** [***2012 USD*** /ton]			
Mean	35,975	45,085	58,401	
	Cost/FFG-7 in 1992 [2012 USD, millions]			
Mean	151	189	245	

This page intentionally left blank.

List of abbreviations/acronyms/initialisms/symbols

ASW	antisubmarine warfare
BLS	(U.S.) Bureau of Labor Statistics
CDF	cumulative distribution function
CDI	Chief of Defence Intelligence
CAF	Canadian Armed Forces
CORA	Centre for Operational Research and Analysis
CPI	Consumer Price Index
DF	degrees of freedom
DHH	Directorate of History and Heritage
DMGOR	Directorate Materiel Group Operational Research
DMPP	Director Materiel Policy and Procedures
DND	Department of National Defence
DRDC	Defence Research and Development Canada
DRENET	Defence Research Establishment Network
DSCA	Defense Security Cooperation Agency
FOC	full operational capability
GBP	Great Britain pound
KS	Kolmogorov-Smirnov
Lasso	least absolute shrinkage and selection operator
MSE	mean of squared errors
NAVSEA	Naval Sea Systems Command
NVR	Naval Vessel Register
O&M	operations and maintenance
OSINT	open source intelligence
PDF	probability density function
PPI	Producer Price Index
RV	random variable
SCN	Shipbuilding and Conversion, Navy (budgetary category)
SE	standard error (of estimated MSE)
URL	Uniform Resource Locator
USD	U.S. dollars
BY	budget year

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence R&D Canada – CORA Dept. of National Defence, MGen G. R. Pearkes Bldg., 101 Colonel By Drive, Ottawa ON K1A 0K2, Canada	2a. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED 2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Sale price as a component of life cycle costing for second hand platforms : Depreciation modelling in the context of the FFG-7 frigate

4. AUTHORS (Last name, followed by initials – ranks, titles, etc. not to be used.)

Ma, F.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information. Include Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
December 2013	81	64

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Memorandum

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence R&D Canada – CORA
Dept. of National Defence, MGen G. R. Pearkes Bldg., 101 Colonel By Drive, Ottawa ON K1A 0K2, Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DRDC CORA TM 2013–227	

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
(   ) Defence departments and defence contractors; further distribution only as approved
(   ) Defence departments and Canadian defence contractors; further distribution only as approved
(   ) Government departments and agencies; further distribution only as approved
(   ) Defence departments; further distribution only as approved
(   ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11)) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

In Q3 of 2012, as a subtask within a Director Materiel Policy and Procedures Major Equipment Procurement Study (New versus Used), Directorate Materiel Group Operational Research (DMGOR) planned a study into the life cycle costing of second hand platforms (vehicles, ships, aircraft). As an initial input into that study, DMGOR undertook the estimation of the depreciation of two classes of frigates, the U.S. FFG-7 and the Dutch Kortenaer, from second hand sales prices. That work is reported here. A database of both frigate classes was developed from open source, but new and used price data were readily available only for the FFG-7s. Depreciation fit an exponential decay model, with an average loss of 8.4%/year and a 68% (± 1 sigma) confidence interval of [5.9%, 11.1%]/year. One variable, vessel age, explained up to 56% of the second used sale price data. The data was not sufficient to extract the portion of depreciation due to aging design/technology rather than physical aging of the platform, but neither did it contradict published cost growth trend of 2%/year for buying new vessels of this type.

Because of the projected growth in real cost for new defence platforms, a thorough analysis of procurement options is becoming increasingly important, including the life cycle costing of buying used. This is an essential part of comparing with buying new that ideally includes platform suitability, proficiency, and timeliness of full operational capability. Existing cost models within DMGOR can potentially be developed for buying used. Exploration of this approach entails examination of disruptive effects due to change of ownership, the availability of data, the effect of aging design/technology on the estimation of depreciation, and the effect of value enhancing accessories such as sensor/weapons systems, accompanying aircraft, and support services. Methods developed for analyzing second hand purchases can potentially be adapted for other major platforms, including other maritime vessels, vehicles, and aircraft. They can also help situate the price range for selling assets, if the Canadian Armed Forces choose to consider this.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Anderson-Darling
categorical variables
confidence bands
cost
depreciation
Doornik-Hansen
FFG-7
frigate
Kolmogorov-Smirnov
Kortenaer
lasso
life cycle cost
linear regression
multiple regression
normal probability plot
normality
Q-Q plot
quantile-quantile plot
regularization
used
warship

DEFENCE DÉFENSE

DRDC CORA

www.drdc-rddc.gc.ca