I am writing in a response to your request to evaluate the dissertation by Alena Skolkova for satisfying requirements for a PhD thesis in economics.

I read the dissertation attentively and with interest. In summary, I think that the dissertation satisfies the requirements for a PhD thesis in economics and should be recommended for a defense. The formal requirement for a dissertation at many PhD programs in economics is that it should consists of three original research papers of a publishable quality. To my opinion, the submitted dissertation does contain three original pieces of research of publishable quality, that satisfy current standards of the field of econometrics. I do support awarding Alena Skolkova the PhD degree in economics.

The rest of my report contains some comments and suggestions, most of them are intended to improve discussions and expositions and are NOT conditions of my recommendation. My recommendation for awarding a degree is unconditional.

(1) This is a general comment for a future direction of improving/extending the paper, it is not expected to be addressed now. I think the first essay would improve significantly from discussions of trade-offs involved in deciding whether to do ridge or OLS. Currently, theoretical model in section 1.3 just demonstrates the benefits without mentioning the costs. Namely, in the current model additional parameters for optimization, λ_1, λ_2 were introduced. The OLS is a special case of ridge when $\lambda_i = 0$. It is expected that allowing to optimize over λ_i would bring some gain, especially when the risk is evaluated at the true optimal values for λ's. The costs though, is that additional parameters (λ's) are introduced, that needed to be estimated from the data, and they potentially bring additional uncertainty. This aspect has not been discussed, and probably a small model like the one in section 1.3 would not be very useful for seeing this aspect. Overall, the question of how to choose λ's is not discussed much, but there may be many interesting aspects, including whether the choice of λ's should depend on the fact that ridge estimators will be used in model averaging, or just choosing an optimal parameter for an individual estimator is enough.

(2) I would prefer to see more details about simulations in section 1.4. Specifically, what is a set of models considered: only the models with first p regressors? or

all models with p some regressors? How many models in total are weighted? It is unclear how the choice of λ's was done. It would be useful to state the details of how the weights were chosen, even if it was stated somewhere else before.

(3) Some discussion of the simulation results would be useful as well. What different choices of α and β mean? Are some cases more sparse/dense than some others? When gains from ridge is more pronounced?

Overall, I enjoyed reading Alena Skolkova's thesis. I congratulate her on successful finishing her dissertation.

Sincerely,
Anna Mikusheva
Professor of Economics, MIT